

A robust ranking method extending ELECTRE III to hierarchy of interacting criteria, imprecise weights and stochastic analysis

Salvatore Corrente^a, José Rui Figueira^b, Salvatore Greco^{a,c}, Roman Słowiński^{d,e}

^a*Department of Economics and Business, University of Catania, Corso Italia, 55, 95129 Catania, Italy*

^b*CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

^c*University of Portsmouth, Portsmouth Business School, Centre of Operations Research and Logistics (CORL), Richmond Building, Portland Street, Portsmouth PO1 3DE, United Kingdom*

^d*Institute of Computing Science, Poznań University of Technology, 60-965 Poznań, Poland*

^e*Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland*

Abstract: A great majority of methods designed for Multiple Criteria Decision Aiding (MCDA) assume that all assessment criteria are considered at the same level, however, decision problems encountered in practice often impose a hierarchical structure of criteria. The hierarchy helps to decompose complex decision problems into smaller and manageable subtasks, and thus, it is very attractive for computational efficiency and explanatory purposes. To handle the hierarchy of criteria in MCDA, a methodology called Multiple Criteria Hierarchy Process (MCHP), has been recently proposed. MCHP permits to consider preference relations with respect to a subset of criteria at any level of the hierarchy. Here, we propose to apply MCHP to the ELECTRE III ranking method adapted to handle three types of interaction effects between criteria: mutual-weakening, mutual-strengthening and antagonistic effect. We also involve in MCHP an imprecise elicitation of criteria weights, generalizing a technique called the SRF method. In order to explore the plurality of rankings obtained by the ELECTRE III method for possible sets of criteria weights, we apply the Stochastic Multiobjective Acceptability Analysis (SMAA) that permits to draw robust conclusions in terms of rankings and preference relations at each level of the hierarchy of criteria. The novelty of the whole methodology consists of a joint consideration of hierarchical assessments of alternatives performances on interacting criteria, imprecise criteria weights, and robust analysis of ranking recommendations resulting from ELECTRE III. An example regarding the multiple criteria ranking of some European universities will show how to apply the proposed methodology on a decision problem.

Keywords: Multiple Criteria Hierarchy Process, ELECTRE III method, SRF method, Stochastic Multiobjective Acceptability Analysis

1. Introduction

Scientific decision aiding is confronted with more and more complex decision problems in various domains of our society, such as environmental management, health care, homeland security, among others. The complexity of these problems is mainly related to the number and the heterogeneity of objectives to be obtained. To deal with these types of problems in an effective way, Multiple Criteria Decision Aiding (MCDA) provides methods that work-out recommendations consistent with

Email addresses: salvatore.corrente@unict.it (Salvatore Corrente), figueira@tecnico.ulisboa.pt (José Rui Figueira), salgreco@unict.it (Salvatore Greco), roman.slowinski@cs.put.poznan.pl (Roman Słowiński)

the preferences of the Decision Maker (DM) [12]. Analysis of the current trends in MCDA leads to the conclusion that the following concerns become more and more relevant in this field:

1. *Consideration of a great number of assessment criteria:* The number of criteria is a factor underlying the complexity of decision problems. The criteria representing different points of view on the quality of alternatives are usually in conflict, because alternatives which are good on some criteria are frequently bad on some other criteria. This makes the search of a compromise among criteria more difficult when the number of criteria is growing [18]. A practical way to cope with the high dimensionality of multiple criteria performances assessment is the hierarchical decomposition of the family of criteria [7]. It permits not only to reduce the number of criteria which contribute to the assessment of alternatives from a specific higher-level point of view, but also to group criteria into logically consistent sub-families. The hierarchical structuring of the family of criteria also facilitates the elicitation of preferences by the DM who is no more obliged to make comparisons of large “vectors” of heterogeneous performances, but can express the preferences with respect to some meaningful subsets of criteria. The DM appreciates, moreover, a final recommendation (in terms of a ranking or a preference relation in the set of alternatives) not only with respect to the whole family of criteria, but also with respect to some sub-criteria located at lower levels of the hierarchy. The above benefits of the hierarchical decomposition, in the sense of preference elicitation and final recommendation, motivated the elaboration of the methodology called Multiple Criteria Hierarchy Process (MCHP) [7].
2. *Consideration of interaction between criteria:* Indeed, it is highly probable that among the many assessment criteria considered in a decision problem, there are interacting criteria in the sense of some mutual-weakening effects, mutual-strengthening effects, or antagonistic effects, as will be explained in subsection 2.2. Different methodologies have been proposed to take into account interaction between criteria, such as non-additive integrals, with the best known Choquet integral for cardinal criteria, and Sugeno integral for ordinal ones [4, 31], some augmented additive value functions [14], and some extended outranking models [9].
3. *Acceptance of intransitivity or incompleteness in the recommended ranking of the alternatives:* In fact, forcing the transitivity and completeness by methods grounded in the Multiple Attribute Value Theory (MAVT) [18], can lead to induce some preferences that are not well founded. Some alternatives may simply be incomparable because of very contrasted performances on the conflicting criteria. The transitivity of an indifference relation is also doubtful, as shown by the famous example of the transitivity of indifference spanned over a series of similar cups of coffee [21]. For these reasons outranking methods, such as ELECTRE and PROMETHEE have been proposed [3, 10] (see [15] for a recent extension of the ELECTRE methods to group decision-making under fuzzy environment). These methods make use of an outranking relation S as a preference model, such that aSb if “alternative a is at least as good as alternative b ”. Thus, if neither aSb nor bSa is true, then a and b are incomparable. The outranking relations are supposed to be only reflexive, i.e., neither complete nor transitive.
4. *Consideration of robustness concerns related to non-univocal values of preference model parameters:* The recommendation supplied by any MCDA method depends on the values assigned to preference model parameters (e.g., weights and veto thresholds); when these values are not univocal (e.g., when many sets of weights and/or thresholds are possible), then it is advisable to identify more and less stable parts of the recommendation. For example, some alternatives may keep the same position in the ranking even if the preference model parameters change in a large interval of possible values, while the position of other alternatives may be much more sensitive to the change of the same model parameters. The range of the changes of a given recommendation is, obviously, an important information for the DM - it gives account of the robustness of this recommendation [25]. In this view, two main approaches have been

developed within MCDA: Stochastic Multiattribute Acceptability Analysis (SMAA) [19] and Robust Ordinal Regression (ROR) [6]. SMAA samples the space of variability of the considered preference model parameters and, in this context, supplies the frequency with which a given alternative is preferred, indifferent or incomparable to another one. ROR, instead, takes into account all sets of preference model parameters compatible with the preference information provided by the DM, and defines the necessary preference relation, in case an alternative is preferred to another one for all the sets of compatible preference model parameters, as well as the possible preference relation, in case an alternative is preferred to another one for at least one set of compatible preference model parameters.

The above four issues are very often present in the same real-life multiple criteria decision problem. There exist methods able to deal with problems where some of these issues are present, however, the MCDA literature does not provide any proof of a joint consideration of all these issues. Consequently, the aim of this paper is to provide a methodology of MCDA that fills this gap. It is worth stressing that elaboration of the comprehensive methodology does not consist of a simple sequential application of the existing methods coping with each particular issue. These issues are interrelated and this requires a careful adaptation of the methods contributing to the comprehensive methodology.

More specifically, we propose to apply MCHP to the ranking method ELECTRE III adapted to handle three types of interaction between criteria: mutual-weakening effect, mutual-strengthening effect, or antagonistic effect. Moreover, we introduce in MCHP an imprecise elicitation of criteria weights, generalizing a technique called the SRF method [11]. Then, to explore the plurality of partial preorders obtained by the ELECTRE III method for possible sets of criteria weights, we apply SMAA that permits to draw robust conclusions in terms of frequency of preference, indifference and incomparability relations at each level of the hierarchy of criteria.

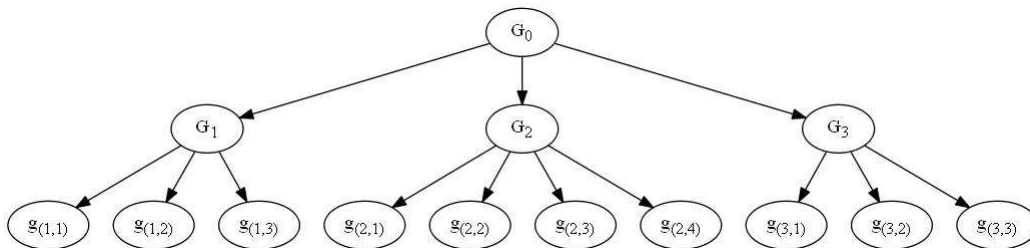
The plan of the article is the following. In Section 2, we recall some basic concepts and methods that contribute to the proposed methodology. This creates a background for next sections where we develop our proposal. A new version of the SRF method handling an imprecise information received from the DM is described in Section 3. In Section 4, we describe the complete methodology, and then, we illustrate its application using a case study presented in Section 5. Section 6 contains conclusions and perspective directions of research.

2. Background

2.1. The Multiple Criteria Hierarchy Process (MCHP)

The MCHP has been introduced in [7] to deal with decision problems where assessment criteria are not considered at the same level but they are organized in a hierarchical way, as shown in Figure 1.

Figure 1: Example of a hierarchy of criteria



To integrate the MCHP with the ELECTRE III method, the following notation will be used:

- G_0 is the root of the hierarchy, and it represents the global objective,
- \mathcal{G} is the set of all criteria in the hierarchy, while $\mathcal{I}_{\mathcal{G}}$ denotes the set of indices of all criteria in the hierarchy,
- EL is the set of indices of all elementary criteria, that are the criteria located at the bottom of the hierarchy,
- $G_{\mathbf{r}}$, where $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, is a generic criterion in the hierarchy, while if $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, then $G_{(\mathbf{r},1)}, \dots, G_{(\mathbf{r},n(\mathbf{r}))}$ are the subcriteria of $G_{\mathbf{r}}$ located at the level immediately below,
- for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $E(G_{\mathbf{r}})$ represents the set of all elementary criteria descending from criterion $G_{\mathbf{r}}$. In the following, a criterion $G_{\mathbf{r}}$, such that $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, will be called *non-elementary criterion*.

2.2. The hierarchical ELECTRE III method with interactions between criteria

In order to apply the ELECTRE III method in conjunction with the MCHP, one needs to get from the DM some preference information in terms of the values of preference model parameters: weights of the elementary criteria, and veto thresholds on these criteria. For each elementary criterion $g_{\mathbf{t}}$, $\mathbf{t} \in EL$, a weight representing its relative importance within the family of elementary criteria is denoted by $w_{\mathbf{t}}$. The weights are supposed to be positive and normalized, that is, $w_{\mathbf{t}} > 0$ for each $\mathbf{t} \in EL$ and $\sum_{\mathbf{t} \in EL} w_{\mathbf{t}} = 1$. Moreover, for each $\mathbf{t} \in EL$ and for each a belonging to the set of alternatives A , the following thresholds should be specified: indifference $q_{\mathbf{t}}(a)$, preference $p_{\mathbf{t}}(a)$, and veto $v_{\mathbf{t}}(a)$ thresholds. The indifference threshold is the maximum difference between the performances of alternatives a and b on $g_{\mathbf{t}}$ compatible with their indifference on $g_{\mathbf{t}}$; the preference threshold is the minimum difference between the performances of a and b on $g_{\mathbf{t}}$ compatible with the preference of one over the other on $g_{\mathbf{t}}$; the veto threshold represents the minimum difference between the performances of b over a on $g_{\mathbf{t}}$ incompatible with the outranking of a over b on any criterion $G_{\mathbf{r}}$ from which $g_{\mathbf{t}}$ is descending, that is, $\mathbf{t} \in E(G_{\mathbf{r}})$. In the following, for the sake of simplicity, but without loss of generality, we shall suppose that the thresholds are not dependent on the performances of the alternatives. Consequently, we shall write $q_{\mathbf{t}}$, $p_{\mathbf{t}}$ and $v_{\mathbf{t}}$ instead of $q_{\mathbf{t}}(a)$, $p_{\mathbf{t}}(a)$ and $v_{\mathbf{t}}(a)$, respectively. ELECTRE III [23] builds a fuzzy outranking relation in the set of alternatives, representing the credibility that an alternative is at least as good as another one. The construction involves concordance and discordance indices concerning the hypothesis that one alternative is at least as good as (outranks) another one. The concordance index represents the strength of the coalition of criteria being in favor of the hypothesis that a outranks b ($a, b \in A$), and the discordance indices represent opposition of criteria being discordant with this hypothesis. In order to perform these tests and calculate the credibility of outranking that aggregates concordance and discordance, one has to calculate for each pair of alternatives $(a, b) \in A \times A$ the following indices [8]:

- the *elementary concordance index*, for each elementary criterion $g_{\mathbf{t}}$,

$$\varphi_{\mathbf{t}}(a, b) = \begin{cases} 1 & \text{if } g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \leq q_{\mathbf{t}}, (aS_{\mathbf{t}}b) \\ \frac{p_{\mathbf{t}} - [g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a)]}{p_{\mathbf{t}} - q_{\mathbf{t}}} & \text{if } q_{\mathbf{t}} < g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < p_{\mathbf{t}}, (bQ_{\mathbf{t}}a) \\ 0 & \text{if } g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq p_{\mathbf{t}}, (bP_{\mathbf{t}}a). \end{cases} \quad (1)$$

- the *elementary discordance index*, for each elementary criterion $g_{\mathbf{t}}$,

$$d_{\mathbf{t}}(a, b) = \begin{cases} 1 & \text{if } g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}}, \\ \frac{[g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a)] - p_{\mathbf{t}}}{v_{\mathbf{t}} - p_{\mathbf{t}}} & \text{if } p_{\mathbf{t}} < g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}, \\ 0 & \text{if } g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \leq p_{\mathbf{t}}, \end{cases} \quad (2)$$

- the *partial concordance index* for each non-elementary criterion $G_{\mathbf{r}}$,

$$C_{\mathbf{r}}(a, b) = \frac{\sum_{\mathbf{t} \in E(G_{\mathbf{r}})} w_{\mathbf{t}} \varphi_{\mathbf{t}}(a, b)}{\sum_{\mathbf{t} \in EL} w_{\mathbf{t}}} \quad (3)$$

- the *partial credibility index* for each non-elementary criterion $G_{\mathbf{r}}$,

$$\sigma_{\mathbf{r}}(a, b) = C_{\mathbf{r}}(a, b) \prod_{\{\mathbf{t} \in E(G_{\mathbf{r}}) : d_{\mathbf{t}}(a, b) > C_{\mathbf{r}}(a, b)\}} \frac{1 - d_{\mathbf{t}}(a, b)}{1 - C_{\mathbf{r}}(a, b)}. \quad (4)$$

In view of the above formulas, observe that in order to avoid that only one immediate subcriterion of $G_{\mathbf{r}}$ decides about outranking ($S_{\mathbf{r}}$) or non-outranking ($S_{\mathbf{r}}^c$) for any pair of alternatives $(a, b) \in A \times A$, even without imposing the veto, it is reasonable to impose a condition that no weight $W_{(\mathbf{r},j)}$ of subcriterion $G_{(\mathbf{r},j)}$, $j = 1, \dots, n(\mathbf{r})$, is greater than the sum of weights of the remaining immediate subcriteria of $G_{\mathbf{r}}$, i.e.,

$$W_{(\mathbf{r},j)} \leq \sum_{\substack{k=1 \\ k \neq j}}^{n(\mathbf{r})} W_{(\mathbf{r},k)}. \quad (5)$$

Otherwise, subcriterion $G_{(\mathbf{r},j)}$ would be a dictator in relation to $G_{\mathbf{r}}$. For this reason, below, we call condition (5) a non-dictatorship condition [26].

In some cases, three types of interactions could be observed between criteria. Given an elementary criterion $g_{\mathbf{t}_1}$ being in favour of the outranking of a over b , then [9]:

- if $g_{\mathbf{t}_2}$ is also in favour of the outranking of a over b , we say that $g_{\mathbf{t}_1}$ and $g_{\mathbf{t}_2}$ present a *mutual-strengthening effect* when the importance assigned to the set composed of the two elementary criteria considered together is greater than the importance assigned to the two elementary criteria considered separately; in this case, a value $w_{\mathbf{t}_1 \mathbf{t}_2} > 0$ has to be added to the sum of the weights $w_{\mathbf{t}_1} + w_{\mathbf{t}_2}$ when computing the concordance index,
- if $g_{\mathbf{t}_2}$ is also in favour of the outranking of a over b , we say that $g_{\mathbf{t}_1}$ and $g_{\mathbf{t}_2}$ present a *mutual-weakening effect* when the importance assigned to the set composed of the two elementary criteria considered together is lower than the importance assigned to the two elementary criteria considered separately; in this case, a value $w_{\mathbf{t}_1 \mathbf{t}_2} < 0$ has to be added to the sum of the weights $w_{\mathbf{t}_1} + w_{\mathbf{t}_2}$ when computing the concordance index,
- if $g_{\mathbf{t}_2}$ is opposing to the outranking of a over b , we say that $g_{\mathbf{t}_2}$ presents an *antagonistic effect* over $g_{\mathbf{t}_1}$ when the importance assigned to $g_{\mathbf{t}_1}$ has to be lowered by the opposition of $g_{\mathbf{t}_2}$. In this case, a value $w'_{\mathbf{t}_1 \mathbf{t}_2} > 0$ has to be subtracted from the weight $w_{\mathbf{t}_1}$ when computing the concordance index.

Let us precise that the contribution given by each criterion to the credibility of the outranking of one alternative over another could not be negative. Therefore, the following net flow condition has to be fulfilled:

Condition 2.1. (see [9])

For all $\mathbf{t}_1 \in EL$,

$$w_{\mathbf{t}_1} - \left[\sum_{\mathbf{t}_2 \in EL: w_{\mathbf{t}_1 \mathbf{t}_2} < 0} |w_{\mathbf{t}_1 \mathbf{t}_2}| + \sum_{\mathbf{t}_3 \in EL} w'_{\mathbf{t}_1 \mathbf{t}_3} \right] \geq 0. \quad (6)$$

Let us denote by $C(bHa)$ the set of all elementary criteria such that bHa , where $H \in \{S, Q, P\}$ and by $\overline{C}(bHa)$ its complement. $\overline{C}(bPa)$ represents the set of elementary criteria that do not oppose too strongly to the outranking of a over b and, consequently, $\overline{C}(bPa) = C(aSb) \cup C(bQa)$. Given a non-elementary criterion $G_{\mathbf{r}}$, and the possible interactions between the elementary criteria, the partial concordance index $C_{\mathbf{r}}(a, b)$ can be rewritten as follows:

$$C_{\mathbf{r}}(a, b) = \frac{1}{W_{\mathbf{r}}(a, b)} \left[\sum_{\mathbf{t}_1 \in \overline{C}(bPa) \cap E(G_{\mathbf{r}})} w_{\mathbf{t}_1} \varphi_{\mathbf{t}_1}(a, b) + \sum_{\mathbf{t}_1, \mathbf{t}_2 \in \overline{C}(bPa) \cap E(G_{\mathbf{r}})} w_{\mathbf{t}_1 \mathbf{t}_2} Z(\varphi_{\mathbf{t}_1}(a, b), \varphi_{\mathbf{t}_2}(a, b)) - \sum_{\substack{\mathbf{t}_1 \in \overline{C}(bPa) \cap E(G_{\mathbf{r}}), \\ \mathbf{t}_2 \in C(bPa) \cap E(G_{\mathbf{r}})}} w'_{\mathbf{t}_1 \mathbf{t}_2} Z(\varphi_{\mathbf{t}_1}(a, b), \varphi_{\mathbf{t}_2}(a, b)) \right] \quad (7)$$

where

$$W_{\mathbf{r}}(a, b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} w_{\mathbf{t}} + \sum_{\mathbf{t}_1, \mathbf{t}_2 \in \overline{C}(bPa) \cap E(G_{\mathbf{r}})} w_{\mathbf{t}_1 \mathbf{t}_2} Z(\varphi_{\mathbf{t}_1}(a, b), \varphi_{\mathbf{t}_2}(a, b)) - \sum_{\substack{\mathbf{t}_1 \in \overline{C}(bPa) \cap E(G_{\mathbf{r}}), \\ \mathbf{t}_2 \in C(bPa) \cap E(G_{\mathbf{r}})}} w'_{\mathbf{t}_1 \mathbf{t}_2} Z(\varphi_{\mathbf{t}_1}(a, b), \varphi_{\mathbf{t}_2}(b, a))$$

and $Z : [0, 1]^2 \rightarrow [0, 1]$ is a non-decreasing function of its arguments, such that $Z(1, x) = Z(x, 1) = x$ for all $x \in [0, 1]$. In the following, we shall consider $Z(x, y) = xy$.

2.2.1. Ascending and Descending distillations

The final partial preorder of alternatives produced by the ELECTRE III method is obtained as the “intersection” of two complete preorders that are the results of the so-called *descending and ascending distillations* [23]. In the descending distillation, one orders the alternatives from the best to the worst, while in the ascending distillation, one orders the alternatives in the opposite way, starting from the worst and finishing with the best.

Given a non-elementary criterion $G_{\mathbf{r}}$, and $\lambda_k \in [0, 1]$, both distillations concern the following fuzzy outranking relation:

$$aS_{\mathbf{r}}^{\lambda_k} b \Leftrightarrow \begin{cases} \sigma_{\mathbf{r}}(a, b) > \lambda_k, \\ \sigma_{\mathbf{r}}(a, b) > \sigma_{\mathbf{r}}(b, a) + s(\sigma_{\mathbf{r}}(a, b)), \end{cases} \quad (8)$$

where $s(\lambda) = \alpha\lambda + \beta$ and, following [23], $\alpha = -0.15$ and $\beta = 0.3$.

For each alternative a , and each non-elementary criterion $G_{\mathbf{r}}$, its λ_k -qualification with respect to $G_{\mathbf{r}}$ is obtained as follows:

$$q_{\mathbf{r}, A}^{\lambda_k}(a) = p_{\mathbf{r}, A}^{\lambda_k}(a) - f_{\mathbf{r}, A}^{\lambda_k}(a) \quad (9)$$

where

- $p_{\mathbf{r}, A}^{\lambda_k}(a) = |\{b \in A : aS_{\mathbf{r}}^{\lambda_k} b\}|$ is the λ_k -power of a ; it is the number of alternatives that are outranked by a on $G_{\mathbf{r}}$,
- $f_{\mathbf{r}, A}^{\lambda_k}(a) = |\{b \in A : bS_{\mathbf{r}}^{\lambda_k} a\}|$ is the λ_k -weakness of a ; it is the number of alternatives outranking a on $G_{\mathbf{r}}$.

Algorithm 1 Descending distillation

Input: Credibility of outranking $\sigma_{\mathbf{r}}(a, b)$ for a given non-elementary criterion $G_{\mathbf{r}}$, and all $a, b \in A$.

Output: Ordered classes of alternatives $\bar{C}_1 \succ \bar{C}_2 \succ \dots \succ \bar{C}_n$, such that $\cup_{i=1}^n \bar{C}_i = A$, where \succ denotes strict preference relation,

1. $n = 0$, $\bar{A}_0 = A$, (\bar{A}_0 contains all alternatives that have not yet been assigned),
2. $\lambda_0 = \max_{a, b \in \bar{A}_n, a \neq b} \{\sigma_{\mathbf{r}}(a, b)\}$,
3. $k = 0$, $D_0 = \bar{A}_n$,
4. $\lambda_{k+1} = \max_{\{(a, b) : \sigma_{\mathbf{r}}(a, b) < \lambda_k - s(\lambda_k), a, b \in D_k\}} \{\sigma_{\mathbf{r}}(a, b)\}$;
 if $\{(a, b) : \sigma_{\mathbf{r}}(a, b) < \lambda_k - s(\lambda_k), a, b \in D_k\} = \emptyset$, **then** $\lambda_{k+1} = 0$,
5. Calculate the λ_{k+1} -qualifications of all the alternatives belonging to D_k with respect to $G_{\mathbf{r}}$,
6. Obtain the maximum λ_{k+1} -qualification $\bar{q}_{\mathbf{r}, D_k} = \max_{a \in D_k} \{q_{\mathbf{r}, D_k}^{\lambda_{k+1}}(a)\}$
7. Build the subset:

$$\bar{D}_{k+1} = \{a \in D_k : q_{\mathbf{r}, D_k}^{\lambda_{k+1}}(a) = \bar{q}_{\mathbf{r}, D_k}\}$$

8. **if** $|\bar{D}_{k+1}| = 1$ or $\lambda_{k+1} = 0$
 go to step 9.
 else
 $k = k + 1$, $D_k = \bar{D}_k$ and go to step 4.
 endif
 9. $\bar{C}_{n+1} = \bar{D}_{k+1}$,
 $\bar{A}_{n+1} = \bar{A}_n \setminus \bar{C}_{n+1}$,
 if $\bar{A}_{n+1} \neq \emptyset$
 $n = n + 1$, go to step 2.
 else
 end of the distillation.
 endif
-

Algorithm 2 Ascending distillation

Input: Credibility of outranking $\sigma_{\mathbf{r}}(a, b)$ for a given non-elementary criterion $G_{\mathbf{r}}$, and all $a, b \in A$.

Output: Ordered classes of alternatives $\underline{C}_1 \succ \underline{C}_2 \succ \dots \succ \underline{C}_n$, such that $\cup_{i=1}^n \underline{C}_i = A$, where \succ denotes strict preference relation,

1. $n = 0$, $\underline{A}_0 = A$, (\overline{A}_0 contains all alternatives that have not yet been assigned),
2. $\lambda_0 = \max_{a, b \in \underline{A}_n, a \neq b} \{\sigma_{\mathbf{r}}(a, b)\}$,
3. $k = 0$, $D_0 = \underline{A}_n$,
4. $\lambda_{k+1} = \max_{\{(a, b): \sigma_{\mathbf{r}}(a, b) < \lambda_k - s(\lambda_k), a, b \in D_k\}} \{\sigma_{\mathbf{r}}(a, b)\}$;
 if $\{(a, b) : \sigma_{\mathbf{r}}(a, b) < \lambda_k - s(\lambda_k), a, b \in D_k\} = \emptyset$, **then** $\lambda_{k+1} = 0$,
5. Calculate the λ_{k+1} -qualifications of all the alternatives belonging to D_k with respect to $G_{\mathbf{r}}$,
6. Obtain the minimum λ_{k+1} -qualification $q_{\mathbf{r}, D_k} = \min_{a \in D_k} \{q_{\mathbf{r}, D_k}^{\lambda_{k+1}}(a)\}$
7. Build the subset:

$$\underline{D}_{k+1} = \{a \in D_k : q_{\mathbf{r}, D_k}^{\lambda_{k+1}}(a) = q_{\mathbf{r}, D_k}\}$$

8. **if** $|\overline{D}_{k+1}| = 1$ or $\lambda_{k+1} = 0$
 go to step 9.
 else
 $k = k + 1$, $D_k = \underline{D}_k$ and go to step 4.
 endif
 9. $\underline{C}_{n+1} = \underline{D}_{k+1}$,
 $\underline{A}_{n+1} = \underline{A}_n \setminus \underline{C}_{n+1}$,
 if $\underline{A}_{n+1} \neq \emptyset$
 $n = n + 1$, go to step 2.
 else
 end of the distillation.
 endif
-

The algorithms for the descending and ascending distillation are described in detail as Algorithm 1 and Algorithm 2, respectively [2].

In each of the distillations the alternatives are ordered in classes, and each class contains at least one alternative. Therefore, the final preorder is obtained as the intersection of the two distillations. Given $a, b \in A$ and a non-elementary criterion $G_{\mathbf{r}}$,

- a is preferred to b on $G_{\mathbf{r}}$ ($aP_{\mathbf{r}}b$), if a belongs to a class not worse than that of b in both distillations, and to a better class for at least one of the two distillations,
- a is indifferent to b on $G_{\mathbf{r}}$ ($aI_{\mathbf{r}}b$), if a and b belong to the same class in both distillations,
- a and b are incomparable on $G_{\mathbf{r}}$ ($aR_{\mathbf{r}}b$), if a belongs to a class better than b in one distillation and worse in the other one.

2.3. The SRF method

As described in subsection 2.2, the application of the ELECTRE III method necessitates the knowledge of weights of the elementary criteria. One can get these weights in two ways: receiving the weights directly from the DM, or using an indirect preference information provided by the DM from which some weights compatible with this information are inferred. In the second way, the Simos method has been proposed in [27, 28], modified later in [11], and implemented in the software called SRF. Hence, we call here the implemented method by the same name - the SRF method.

In the following, we describe the application of the SRF method to the set composed of elementary criteria only. Let us remind that the application of the SRF method to the case of criteria structured in a hierarchical way has been introduced in [5].

In the SRF method, the DM is asked to rank the elementary criteria with respect to their relative importance from the least important, belonging to the subset $L_1 \subseteq EL$, to the most important, belonging to the subset $L_v \subseteq EL$. Moreover, the DM can insert some blank cards between two successive subsets of criteria L_s and L_{s+1} to increase the difference of importance between elementary criteria in L_s and in L_{s+1} . The number of blank cards included between subsets of elementary criteria L_s and L_{s+1} is denoted by e_s . Finally, the DM is asked to provide a ratio z between the importance of elementary criteria in L_v and the importance of elementary criteria in L_1 .

From a formal point of view, following [5], the non-normalized weight of elementary criterion $g_{\mathbf{t}}$ is computed as

$$w'_{\mathbf{t}} = 1 + \frac{(z - 1) \left[l(\mathbf{t}) - 1 + \sum_{s=1}^{l(\mathbf{t})-1} e_s \right]}{v - 1 + \sum_{s=1}^{v-1} e_s}$$

where $l(\mathbf{t})$ represents the rank to which elementary criterion $g_{\mathbf{t}}$ belongs, that is $\mathbf{t} \in L_{l(\mathbf{t})}$. Consequently, the normalized weight $w_{\mathbf{t}}$ assigned to each elementary criterion $\mathbf{t} \in EL$ is obtained as follows:

$$w_{\mathbf{t}} = \frac{w'_{\mathbf{t}}}{\sum_{\mathbf{t} \in EL} w'_{\mathbf{t}}}$$

2.4. The Stochastic Multiobjective Acceptability Analysis (SMAA)

The SMAA methodology has been presented for the first time in [19] in order to deal with decision problems where evaluations of alternatives on criteria, and some preference model parameters are not precisely known. SMAA can be considered as a way of handling robustness concerns related to

preference model parameters and performances of alternatives on the considered criteria. Below, we remind the main features of SMAA without assuming any specific preference model (for a survey of the SMAA methods see [32], while for a real-world application of SMAA see [33]). In SMAA two probability distributions, denoted by f_W and f_χ , are defined on W and the performance space $\chi = [g_j(a) \mid j = 1, \dots, m, \text{ and } a \in A]$. SMAA is based on sampling of both weight and performance spaces. Depending on the chosen preference model various indices can be used to summarize the sampling results. For preference model being a value function (providing a complete ranking of the alternatives), the *rank acceptability index* $b^l(a)$ and the *central weight vector* $w^c(a)$ can be computed for each $a \in A$ and for each rank position $l = 1, \dots, n$. $b^l(a)$ gives the frequency with which alternative a takes the l -th position in the final ranking, while $w^c(a)$ specifies the average weight vector giving to alternative a the best position in the final ranking.

For a preference model being a value function or an outranking relation, the *pairwise winning index* $Pre(a, c)$ gives the frequency with which alternative a is preferred to alternative c [20]

$$Pre(a, c) = \int_{w \in W} f_W(w) \int_{\xi \in \chi: a \succ^M c} f_\chi(\xi) d\xi dw$$

where $a \succ^M c$ is the preference relation provided by the model compatible with the DM's preferences.

3. The imprecise SRF method

As already observed in [29], the SRF method produces a single set of criteria weights while, in general, there exists an infinity of the sets of weights compatible with the preference information provided by the DM. For this reason, the authors of [29] proposed different ways of handling the plurality of sets of the weights compatible with the preference information provided by the DM. In the present paper, our aim is quite different, since we propose a new variant of the SRF method that takes into account a different type of imprecision. This imprecision is not only related to the plurality of sets of weights compatible with the preference information provided by the DM but it is also related to the imprecision of this information.

As explained in subsection 2.3, in order to apply the SRF method, the DM is asked to:

- 1) rank the criteria from the least important to the most important with the possibility of ex aequo between them,
- 2) put some blank cards between two successive subsets of criteria to increase the difference of importance between criteria in these two subsets,
- 3) define the ratio z between the importance of the most important subset of criteria and the importance of the least important one.

We suppose that the imprecision of the information provided by the DM regards points 2) and 3). Indeed, in our opinion, when applying the SRF method, it is quite reasonable to suppose that the DM may provide, on one hand, an imprecise number of blank cards between two successive subsets of criteria and, on the other hand, an imprecise importance ratio z .

In the following, first we describe the imprecise SRF method on a set of generic criteria $\{g_1, \dots, g_m\}$, after we extend the method to the case of a set of weights organized in a hierarchy.

Let us suppose that the DM claims that criteria in L_{s+1} are more important than criteria in L_s but, she is not sure about the number of blank cards that should be included between the two subsets of criteria, that is, about the value of e_s . Consequently, she states that $e_s \in [low_s, upp_s]$ where, $low_s \leq upp_s$. In this way, denoting the weights of criteria in L_s and L_{s+1} by w_{L_s} and $w_{L_{s+1}}$,

respectively, and denoting by C the importance of a blank card included between L_s and L_{s+1} , the preference information provided by the DM can be translated to the following pair of constraints:

$$\begin{cases} w_{L_{s+1}} \geq w_{L_s} + (low_s + 1) \cdot C, \\ w_{L_{s+1}} \leq w_{L_s} + (upp_s + 1) \cdot C, \\ C > 0. \end{cases}$$

Analogously, it is possible that, the DM is unsure about the ratio z between the criteria in L_v and the criteria in L_1 . For this reason, she may wish to provide an interval of possible values to which z should belong. This means that $z \in [z_{low}, z_{upp}]$. Denoting by w_{L_1} and w_{L_v} the weight of a criterion belonging to L_1 and L_v , respectively, this piece of preference information can be easily translated to the constraints

$$z_{low} \leq \frac{w_{L_v}}{w_{L_1}} \leq z_{upp} \Leftrightarrow z_{low} \cdot w_{L_1} \leq w_{L_v} \leq z_{upp} \cdot w_{L_1}.$$

Therefore, the set of constraints translating the above preference information provided by the DM is the following:

$$\left\{ \begin{array}{l} w_{L_{s+1}} \geq w_{L_s} + (low_s + 1) \cdot C, \\ w_{L_{s+1}} \leq w_{L_s} + (upp_s + 1) \cdot C, \\ C > 0, \end{array} \right\} \text{ for all } s = 1, \dots, v-1, \quad (10)$$

$$\left\{ \begin{array}{l} z_{low} \cdot w_{L_1} - w_{L_v} \leq 0, \\ w_{L_v} - z_{upp} \cdot w_{L_1} \leq 0, \\ w_{L_1} > 0 \end{array} \right.$$

while to normalize the weights of criteria, the following constraint should also be included:

$$\sum_{i=1}^v |L_i| \cdot w_{L_i} = 1.$$

If the imprecise SRF method is applied to a flat structure of criteria, the non-dictatorship condition (5) should be appended to the constraints (10) if L_v would be composed of only one criterion, and only for this criterion. This is because only if L_v is a singleton, this criterion could become a dictator. Therefore, the constraint that should be appended to (10) in this case is the following:

$$w_{L_v} \leq \sum_{s=1}^{v-1} w_{L_s} \cdot |L_s|.$$

Let us observe that the DM can consider the whole set \mathbf{W} of weights satisfying the constraints of the imprecise SRF method, but she can also select one representative set of weights, for example the barycenter of \mathbf{W} .

Example 3.1. *Let us suppose that the DM considers 8 criteria $\{g_1, \dots, g_8\}$ and that she provides the following preference information:*

- *the criteria are ordered from the least important to the most important ones in the following subsets: $L_1 = \{g_6\}$, $L_2 = \{g_5, g_7\}$, $L_3 = \{g_1, g_3\}$, and $L_4 = \{g_2, g_4\}$,*
- *the number of blank cards between L_1 and L_2 should belong to $[1, 3]$, there is one blank card between L_2 and L_3 and the number of blank cards between L_3 and L_4 should belong to $[2, 4]$,*
- *the importance ratio between criteria in L_4 and the criteria in L_1 should belong to the interval $[2, 4]$.*

According to (10), the preference information provided by the DM is translated to the following set of constraints:

$$\left\{ \begin{array}{l} w_{L_1} + 2C \leq w_{L_2} \leq w_{L_1} + 4C, \\ w_{L_3} = w_{L_2} + 2C, \\ w_{L_3} + 3C \leq w_{L_4} \leq w_{L_3} + 5C, \\ 2 \cdot w_{L_1} \leq w_{L_4} \leq 4w_{L_1}, \\ C > 0, \\ w_{L_1} > 0, \\ w_{L_1} + 2 \cdot w_{L_2} + 2 \cdot w_{L_3} + 2 \cdot w_{L_4} = 1. \end{array} \right.$$

where $w_6 = w_{L_1}$, $w_5 = w_7 = w_{L_2}$, $w_1 = w_3 = w_{L_3}$, and $w_2 = w_4 = w_{L_4}$. Whichever vector of weights $(w_{L_1}, w_{L_2}, w_{L_3}, w_{L_4})$ satisfying the above set of constraints is therefore compatible with the preference information provided by the DM. For example, a feasible solution is $w_{L_1} = 0.06$, $w_{L_2} = 0.1$, $w_{L_3} = 0.15$ and $w_{L_4} = 0.22$. Let us underline that, as explained above, in this case it is not necessary to take into account the non-dictatorship condition since two criteria are in the most important subset.

In order to adapt the proposed imprecise SRF method to the case of criteria structured in a hierarchical way, the DM is therefore asked to apply the imprecise SRF method on each subset of criteria $\{G_{(r,1)}, \dots, G_{(r,n(r))}\}$ composed of the immediate subcriteria of the non-elementary criterion G_r . It will be enough to perform the following replacements in the notation of the imprecise SRF method described above:

- L_1, \dots, L_v with L_1^r, \dots, L_v^r ,
- w_{L_1}, \dots, w_{L_v} with $W_{L_1}^r, \dots, W_{L_v}^r$,
- e_s with e_s^r ,
- $[low_s, upp_s]$ with $[low_s^r, upp_s^r]$,
- z with z^r ,
- $[z_{low}, z_{upp}]$ with $[z_{low}^r, z_{upp}^r]$.

Therefore, in applying the imprecise SRF method to the set of criteria $\{G_{(r,1)}, \dots, G_{(r,n(r))}\}$, the set of constraints in (10) will be replaced by the following one:

$$E_r \left\{ \begin{array}{l} W_{L_{s+1}^r} \geq W_{L_s^r} + (low_s^r + 1) \cdot C_r, \\ W_{L_{s+1}^r} \leq W_{L_s^r} + (upp_s^r + 1) \cdot C_r, \\ C_r > 0, \\ z_{low}^r \cdot W_{L_1^r} - W_{L_v^r} \leq 0, \\ W_{L_v^r} - z_{upp}^r \cdot W_{L_1^r} \leq 0, \\ W_{L_1^r} > 0. \end{array} \right\} \text{ for all } s = 1, \dots, v-1, \quad (11)$$

Two more points should be underlined:

1. According to [5], the weight W_r of each non-elementary criterion G_r , can be obtained as the sum of the weights of the elementary criteria descending from G_r , that is

$$W_r = \sum_{t \in E(G_r)} w_t. \quad (12)$$

Consequently, the constraints translating the preference information provided by the DM during the application of the imprecise SRF method on the subset of criteria $\{G_{(r,1)}, \dots, G_{(r,n(r))}\}$, can be expressed in terms of the weights of the elementary criteria only.

2. The value of C_r appearing in the set of constraints (11) and representing the importance of a blank card inserted between two successive subsets of criteria, is not the same whichever criterion in the hierarchy is considered. For example, if the criteria are structured in a hierarchical way as in Figure 1, then a blank card inserted between criteria G_1 and G_2 is not necessarily as important as a blank card inserted between elementary criteria $g_{(2,3)}$ and $g_{(2,2)}$ at the bottom level of the hierarchy.

Example 3.2. *Let us suppose that the DM is interested in obtaining the weights of criteria structured in a hierarchical way as shown in Figure 1. For this reason, she applies the imprecise SRF method at first on the set of criteria $\{G_1, G_2, G_3\}$, and then, on the sets of criteria $\{g_{(1,1)}, g_{(1,2)}, g_{(1,3)}\}$, $\{g_{(2,1)}, g_{(2,2)}, g_{(2,3)}, g_{(2,4)}\}$ and $\{g_{(3,1)}, g_{(3,2)}, g_{(3,3)}\}$, separately. In the following, we shall show in detail the application of the imprecise SRF method to the sets of criteria $\{G_1, G_2, G_3\}$ and $\{g_{(3,1)}, g_{(3,2)}, g_{(3,3)}\}$ only.*

- *Regarding the first level criteria, the DM states that $G_2 = L_1^0$ is less important than $G_1 = L_2^0$ that is less important than $G_3 = L_3^0$. She says that the number of blank cards e_1^0 between G_2 and G_1 belongs to the interval $[2, 3] = [low_1^0, upp_1^0]$, while the number of blank cards e_2^0 between G_1 and G_3 belongs to the interval $[1, 2] = [low_2^0, upp_2^0]$. Moreover, the importance ratio z^0 between G_3 and G_2 belongs to the interval $[1, 3] = [z_{low}^0, z_{upp}^0]$. Consequently, denoting by C_0 the importance of a blank card inserted between two successive subsets of criteria, the set of constraints translating these preferences is the following:*

$$E_0 \left\{ \begin{array}{l} W_{L_1^0} + 3C_0 \leq W_{L_2^0} \leq W_{L_1^0} + 4C_0, \\ W_{L_2^0} + 2C_0 \leq W_{L_3^0} \leq W_{L_2^0} + 3C_0, \\ W_{L_1^0} \leq W_{L_3^0} \leq 3W_{L_1^0}, \\ W_{L_1^0} > 0, \\ C_0 > 0. \end{array} \right. \rightarrow \left\{ \begin{array}{l} \sum_{t \in E(G_2)} w_t + 3C_0 \leq \sum_{t \in E(G_1)} w_t \leq \sum_{t \in E(G_2)} w_t + 4C_0, \\ \sum_{t \in E(G_1)} w_t + 2C_0 \leq \sum_{t \in E(G_3)} w_t \leq \sum_{t \in E(G_1)} w_t + 3C_0, \\ \sum_{t \in E(G_2)} w_t \leq \sum_{t \in E(G_3)} w_t \leq 3 \sum_{t \in E(G_2)} w_t, \\ \sum_{t \in E(G_2)} w_t > 0, \\ C_0 > 0. \end{array} \right. \quad (13)$$

- *Regarding the subset of elementary criteria $\{g_{(3,1)}, g_{(3,2)}, g_{(3,3)}\}$, let us suppose that the DM states that $g_{(3,1)} = L_1^3$ is less important than $g_{(3,3)} = L_2^3$ that, in turn, is less important than $g_{(3,2)} = L_3^3$. She says that the number of blank cards e_1^3 between $g_{(3,1)}$ and $g_{(3,3)}$ should belong to the interval $[1, 2] = [low_1^3, upp_1^3]$, while the number of blank cards e_2^3 between $g_{(3,3)}$ and $g_{(3,2)}$ should belong to the interval $[2, 3] = [low_2^3, upp_2^3]$. Moreover, the importance ratio z^3 between $g_{(3,2)}$ and $g_{(3,1)}$ should belong to the interval $[1, 4] = [z_{low}^3, z_{upp}^3]$. As a consequence, denoting by C_3 the importance of a blank card inserted between two successive subsets of criteria, this*

preference information can be translated into the following set of constraints:

$$E_3 \left\{ \begin{array}{l} W_{L_1^3} + 2C_3 \leq W_{L_2^3} \leq W_{L_1^3} + 3C_3, \\ W_{L_2^3} + 3C_3 \leq W_{L_3^3} \leq W_{L_2^3} + 4C_3, \\ W_{L_1^3} \leq W_{L_3^3} \leq 4W_{L_1^3}, \\ W_{L_1^3} > 0, \\ C_3 > 0. \end{array} \right. \rightarrow \left\{ \begin{array}{l} w_{(\mathbf{3},1)} + 2C_3 \leq w_{(\mathbf{3},3)} \leq w_{(\mathbf{3},1)} + 3C_3, \\ w_{(\mathbf{3},3)} + 3C_3 \leq w_{(\mathbf{3},2)} \leq w_{(\mathbf{3},3)} + 4C_3, \\ w_{(\mathbf{3},1)} \leq w_{(\mathbf{3},2)} \leq 4w_{(\mathbf{3},1)}, \\ w_{(\mathbf{3},1)} > 0, \\ C_3 > 0. \end{array} \right. \quad (14)$$

In order to adapt the condition of non-dictatorship (5) when using the imprecise SRF method to criteria structured hierarchically, we have to take into account two special cases:

case 1) No information is provided by the DM about relative importance of criteria immediately descending from non-elementary criterion $G_{\mathbf{r}}$. In this case, for each $j = 1, \dots, n(\mathbf{r})$, the following constraint should be added:

$$W_{(\mathbf{r},j)} \leq \sum_{\substack{k=1 \\ k \neq j}}^{n(\mathbf{r})} W_{(\mathbf{r},k)};$$

case 2) The DM provided information about relative importance of criteria immediately descending from non-elementary criterion $G_{\mathbf{r}}$. In this case, the non-dictatorship condition should be included only if a single criterion has been indicated as the most important among the immediate descending of $G_{\mathbf{r}}$. Then, the non-dictatorship condition is the following:

$$ND_{\mathbf{r}} \left\{ \begin{array}{l} W_{L_v^{\mathbf{r}}} \leq \sum_{s=1}^{v-1} W_{L_{s-1}^{\mathbf{r}}} \cdot |L_{s-1}^{\mathbf{r}}|, \\ W_{L_s^{\mathbf{r}}} = W_{(\mathbf{r},j)}, \text{ for all } s = 1, \dots, v-1, \text{ and for all } G_{(\mathbf{r},j)} \in L_s^{\mathbf{r}}. \end{array} \right.$$

As in this paper we are considering a hierarchy of interacting criteria, the complete set of constraints defining the weights of the elementary criteria and the coefficients representing the possible interactions between elementary criteria is the following:

$$E \left\{ \begin{array}{l} [1] \cup_{\mathbf{r} \in \mathcal{I}_G \setminus EL} \{E_{\mathbf{r}} \cup ND_{\mathbf{r}}\}, \\ [2] \left\{ \begin{array}{l} w_{\mathbf{t}_1, \mathbf{t}_2} > 0 \text{ if elementary criteria } g_{\mathbf{t}_1} \text{ and } g_{\mathbf{t}_2} \text{ present a mutual-strengthening effect,} \\ w_{\mathbf{t}_1, \mathbf{t}_2} < 0 \text{ if } g_{\mathbf{t}_1} \text{ and } g_{\mathbf{t}_2} \text{ present a mutual-weakening effect,} \\ w'_{\mathbf{t}_1, \mathbf{t}_2} > 0 \text{ if } g_{\mathbf{t}_2} \text{ presents an antagonistic effect over } g_{\mathbf{t}_1}, \end{array} \right. \\ [3] w_{\mathbf{t}_1} - \left[\sum_{\mathbf{t}_2 \in EL: w_{\mathbf{t}_1, \mathbf{t}_2} < 0} |w_{\mathbf{t}_1, \mathbf{t}_2}| + \sum_{\mathbf{t}_3 \in EL} w'_{\mathbf{t}_1, \mathbf{t}_3} \right] \geq 0, \text{ for all } \mathbf{t}_1 \in EL, \\ [4] \sum_{\mathbf{t} \in EL} w_{\mathbf{t}} = 1, \\ [5] \sum_{\substack{\{\mathbf{t}_1, \mathbf{t}_2\} \subseteq EL: \\ w_{\mathbf{t}_1, \mathbf{t}_2} > 0}} w_{\mathbf{t}_1, \mathbf{t}_2} \leq \sum_{\mathbf{t} \in EL} w_{\mathbf{t}}. \end{array} \right.$$

In E , constraints [1] translate the information provided by the DM during the application of the imprecise SRF method on each subset of criteria $\{G_{(\mathbf{r},1)}, \dots, G_{(\mathbf{r},n(\mathbf{r}))}\}$ and include possibly the non-dictatorship condition imposed for the most important criterion immediately descending from $G_{\mathbf{r}}$.

Constraints [2] translate the preference information provided by the DM in terms of possible interactions between elementary criteria, while constraints [3]-[5] are technical constraints. In particular, constraints [3] ensure that the contribution given by each elementary criterion to the credibility of outranking is non-negative; constraint [4] is a normalization constraint, while constraint [5] is introduced to limit the importance assigned to the mutual-strengthening effects. Indeed, in this way, the sum of the coefficients representing the mutual-strengthening effects between the elementary criteria cannot be more important than the unanimity of the elementary criteria.

Let us underline that the methodology does not exclude some additional constraints on the weights if the DM is willing to express them. For example, independently from the application of the SRF method she could claim that a criterion is more important than the sum of others, or that the weight assigned to a criterion has not to be lower than a certain threshold. In those cases, the corresponding constraints have to be added to the set E .

4. The proposed methodology

The set of constraints E defines a space composed of an infinite number of vectors of weights and interaction coefficients compatible with the preference information provided by the DM. In the following, we call such a vector a *compatible vector*. For existence of many compatible vectors, we propose to apply the SMAA methodology. SMAA permits to get robust recommendations by exploring the different partial rankings resulting from the application of the hierarchical ELECTRE III method with interactions on a large sample of compatible vectors.

The flow chart describing the proposed method is shown in Figure 2 and it is composed of the following steps:

Step 0 Fix the hierarchical structure of the set of criteria using the MCHP, distinguishing the subsets of elementary criteria, and higher level criteria, up to the root criterion.

Step 1 The DM, or the analyst representing the DM, is asked to provide different types of preference information:

- 1.1 Indifference, preference and veto thresholds for each elementary criterion.
- 1.2 Information about possible mutual-strengthening and mutual-weakening effects, as well as about the possible antagonistic effects between some elementary criteria.
- 1.3 Information required by the imprecise SRF method for each subset of criteria $\{G_{(r,1)}, \dots, G_{(r,n(r))}\}$ and each non-elementary criterion G_r .

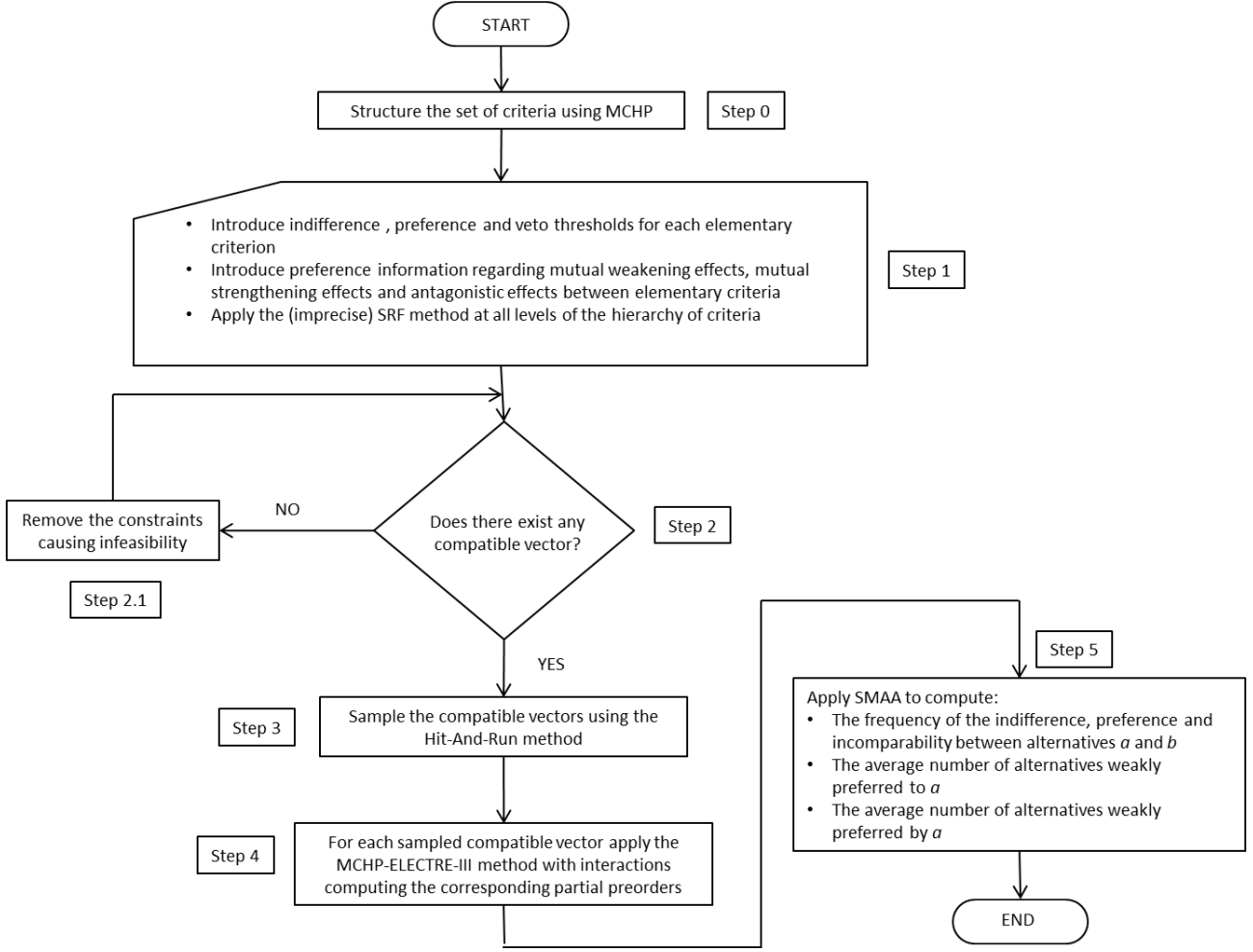
Step 2 The analyst has to check if there exists at least one compatible vector. From a computational point of view, this means that one has to solve the following LP problem:

$$\varepsilon^* = \max \varepsilon, \quad \text{s.t.} \quad E' = [1'] \cup [2'] \cup [3'] \cup [4] \cup [5] \quad (15)$$

where the sets of constraints $[1']$, $[2']$, and $[3']$ are obtained from [1], [2], and [3] by transforming strict inequalities into weak inequalities using an auxiliary variable ε (for example, $C_0 > 0$ is transformed into $C_0 \geq \varepsilon$).

If the set of constraints E' is feasible and $\varepsilon^* > 0$, then there exists at least one compatible vector. In the opposite case (E' is infeasible or $\varepsilon^* \leq 0$), there is some inconsistency in the preference information provided by the DM that can be identified by using some of the techniques proposed in [22] (Step 2.1). Once this inconsistency is removed, one can pass to the next step.

Figure 2: Flow chart of the proposed methodology



Step 3 As there may exist many compatible vectors, we suggest to sample them. Since set E' is composed of linear constraints and they define a convex space, one can use the Hit-And-Run (HAR) method proposed in [30] and recently revisited in [34] and [35].

Step 4 For each sampled compatible vector and for each non-elementary criterion $G_{\mathbf{r}}$, one performs the descending and the ascending distillations of the credibility of outranking, giving descending and ascending complete preorders, respectively. Their intersection gives a partial preorder of the alternatives. Consequently, as observed in Section 2.2.1, the preference ($P_{\mathbf{r}}$), indifference ($I_{\mathbf{r}}$) and incomparability ($R_{\mathbf{r}}$) relations are known for each compatible vector and for each non-elementary criterion $G_{\mathbf{r}}$.

Aggregating the preference and the indifference relations, we also know that a is weakly preferred to b on $G_{\mathbf{r}}$, and we write $aWP_{\mathbf{r}}b$, if a is preferred to b , or a is indifferent to b on $G_{\mathbf{r}}$. Therefore, $WP_{\mathbf{r}} = P_{\mathbf{r}} \cup I_{\mathbf{r}}$.

Step 5 The SMAA methodology is applied to summarize the results of the previous step and to get a robust recommendation. SMAA provides three different types of information:

- 5.1 For each pair of alternatives $(a, b) \in A \times A$ and for each non-elementary criterion $G_{\mathbf{r}}$:
- $Pre_{\mathbf{r}}(a, b)$, i.e., the frequency with which a is preferred to b on $G_{\mathbf{r}}$,
 - $Ind_{\mathbf{r}}(a, b)$, i.e., the frequency with which a is indifferent to b on $G_{\mathbf{r}}$,

- $Inc_{\mathbf{r}}(a, b)$, i.e., the frequency with which a and b are incomparable on $G_{\mathbf{r}}$,
- $Weak_{\mathbf{r}}(a, b)$, i.e., the frequency with which a is weakly preferred to b on $G_{\mathbf{r}}$. Observe that, $Weak_{\mathbf{r}}(a, b) = Pre_{\mathbf{r}}(a, b) + Ind_{\mathbf{r}}(a, b)$ for all $(a, b) \in A \times A$ and for each $G_{\mathbf{r}}$. Moreover, obviously, $Weak_{\mathbf{r}}(a, b) + Inc_{\mathbf{r}}(a, b) + Pref_{\mathbf{r}}(b, a) = 100\%$ for all $(a, b) \in A \times A$ and for each $G_{\mathbf{r}}$.

5.2 For each non-elementary criterion $G_{\mathbf{r}}$ and for each alternative a :

- the average number of alternatives $b \in A$ such that $aWP_{\mathbf{r}}b$; it is the mean number of alternatives to which a is weakly preferred on $G_{\mathbf{r}}$ for all sampled compatible vectors,
- the average number of alternatives $b \in A$ such that $bWP_{\mathbf{r}}a$; it is the mean number of alternatives weakly preferred to a on $G_{\mathbf{r}}$ for all sampled compatible vectors.

5.3 For each non-elementary criterion $G_{\mathbf{r}}$, as well as for the root criterion $G_{\mathbf{0}}$, the whole set of partial preorders supplied by the ELECTRE III exploitation procedure with their related frequencies and the barycenter of the sets of preference parameters originating them.

The information got by the frequencies shown in step 5.1 are linked to the results that one could obtain from the application of the ROR (see [13] for the first publication on ROR and [6] for a survey on ROR from a machine learning perspective). Analogously to the SMAA methodology, the ROR explores the whole space of compatible vectors by building two *necessary* and two *possible* outranking relations for each pair of alternatives $a, b \in A$ and for each non-elementary criterion $G_{\mathbf{r}}$:

- a is weakly necessarily preferred to b on $G_{\mathbf{r}}$, which is denoted by $aWP_{\mathbf{r}}^N b$, if a is weakly preferred to b on $G_{\mathbf{r}}$ for all compatible vectors,
- a is weakly possibly preferred to b on $G_{\mathbf{r}}$, which is denoted by $aWP_{\mathbf{r}}^P b$, if a is weakly preferred to b on $G_{\mathbf{r}}$ for at least one compatible vector,
- a is necessarily not weakly preferred to b on $G_{\mathbf{r}}$, which is denoted by $aWP_{\mathbf{r}}^{cN} b$, if a is not weakly preferred to b on $G_{\mathbf{r}}$ for all compatible vectors,
- a is possibly not weakly preferred to b on $G_{\mathbf{r}}$, which is denoted by $aWP_{\mathbf{r}}^{cP} b$, if a is not weakly preferred to b on $G_{\mathbf{r}}$ for at least one compatible vector.

Putting together the results of SMAA and ROR, one can observe what follows for all $a, b \in A$ and for each non-elementary criterion $G_{\mathbf{r}}$:

- if $Weak_{\mathbf{r}}(a, b) > 0\%$, then $aWP_{\mathbf{r}}^P b$,
- if $aWP_{\mathbf{r}}^N b$ then $Weak_{\mathbf{r}}(a, b) = 100\%$,
- if $aWP_{\mathbf{r}}^{cN} b$, then $Weak_{\mathbf{r}}(a, b) = 0\%$,
- if $Weak_{\mathbf{r}}(a, b) < 100\%$, then $aWP_{\mathbf{r}}^{cP} b$.

Thus, especially if the number of sampled extractions of preference parameters is statistically significant, one can approximate the relations $WP_{\mathbf{r}}^P$, $WP_{\mathbf{r}}^N$, $WP_{\mathbf{r}}^{cN}$, and $WP_{\mathbf{r}}^{cP}$ as follows: for all $a, b \in A$:

- $aWP_{\mathbf{r}}^P b$ iff $Weak_{\mathbf{r}}(a, b) > 0\%$,
- $aWP_{\mathbf{r}}^N b$ iff $Weak_{\mathbf{r}}(a, b) = 100\%$,
- $aWP_{\mathbf{r}}^{cN} b$ iff $Weak_{\mathbf{r}}(a, b) = 0\%$,
- $aWP_{\mathbf{r}}^{cP} b$ iff $Weak_{\mathbf{r}}(a, b) < 100\%$.

The link between SMAA and ROR for ranking and sorting decisions has been investigated in [16] and [17], respectively.

Let us emphasize the innovative steps of the above procedure:

- Step 1.3: the new version of the SRF method permits to take into account imprecise preference information concerning the weights,
- Step 2: the LP problem (15) takes into account the constraints following from imprecise assessment of weights,
- Step 4: the ELECTRE III procedure applied to hierarchically structured interacting criteria is run for each sampled compatible vector, together with distillations.
- Step 5: adaptation of SMAA to preorder which is not necessarily complete.

5. Case study: university ranking

In this section, we show how to apply the proposed methodology to a decision making problem regarding the university ranking. Fourteen European universities, assessed on ten elementary criteria located at the bottom of a hierarchy tree shown in Figure 3, have been extracted from database [1]. Performances of the universities on the considered elementary criteria in the 1-5 scale are shown in Table 1.

Figure 3: Hierarchical structure of criteria considered in the case study

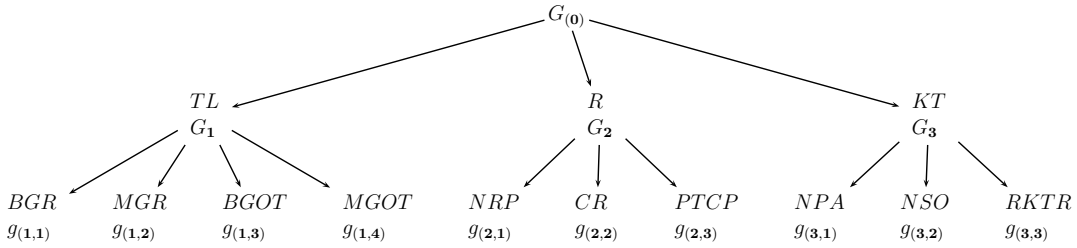


Table 1: Evaluations of the universities at hand on the considered elementary criteria

University	Country	$G_{(0)}$									
		TL ($G_{(1)}$)				R ($G_{(2)}$)			KT ($G_{(3)}$)		
		BGR ($g_{(1,1)}$)	MGR ($g_{(1,2)}$)	BGOT ($g_{(1,3)}$)	MGOT ($g_{(1,4)}$)	NRP ($g_{(2,1)}$)	CR ($g_{(2,2)}$)	PTCP ($g_{(2,3)}$)	NPA ($g_{(3,1)}$)	NSO ($g_{(3,2)}$)	RKTR ($g_{(3,3)}$)
U Bergen (U_1)	Norway	2	4	3	3	5	4	4	4	2	3
Polytech. Inst. Bragança (U_2)	Portugal	4	2	2	3	2	4	4	1	5	3
Dublin Inst. Tech (U_3)	Ireland	5	2	4	4	2	5	5	3	4	2
European U Cyprus (U_4)	Cyprus	3	3	2	2	2	5	4	1	1	4
U Graz (U_5)	Austria	2	3	2	2	3	5	5	2	2	5
Lappeenranta U Tech (U_6)	Finland	5	4	2	3	5	3	3	1	5	4
U Lisbon (U_7)	Portugal	5	2	3	3	5	4	4	2	2	2
U Magdeburg (U_8)	Germany	3	4	2	2	5	3	3	3	1	5
U Montpellier 1 (U_9)	France	3	2	5	5	3	4	5	3	4	3
MEPhI Moscow (U_{10})	Russia	4	2	3	5	5	2	2	1	2	5
UAS Osnabrück (U_{11})	Germany	4	4	3	2	2	4	5	2	4	2
U Pisa (U_{12})	Italy	3	4	2	2	4	4	3	3	2	4
Semmelweis U (U_{13})	Hungary	3	2	4	4	5	3	2	3	2	2
U Thessaly (U_{14})	Greece	4	4	2	2	4	3	2	1	3	4

In Table 2, we provide a description of the ten elementary criteria. As one can see from Figure 3, Teaching and Learning (TL), Research (R) and Knowledge Transfer (KT) are macro-criteria. Masters Graduation Rate (MGR), Masters Graduating on Time (MGT), Bachelor Graduation Rate (BGR) and Bachelor Graduating on Time (BGT) are elementary criteria descending from Teaching and Learning; Number of Research Publications (NRP), Citation Rate (CR) and Proportion of Top Cited Publications (PTCP) are elementary criteria descending from Research, while Number

Table 2: Description of the elementary criteria

Elementary criterion	Description
Bachelor Graduation Rate (BGR)	The percentage of new entrants that successfully completed their bachelor programs
Masters Graduation Rate (MGR)	The percentage of new entrants that successfully completed their master programs
Bachelor Graduating on Time (BGOT)	The percentage of graduates that graduated within the time expected (normative time) for their bachelor programs
Masters Graduating on Time (MGOT)	The percentage of graduates that graduated within the time expected (normative time) for their masters programs
Number of Research Publications (NRP)	The number of research publications indexed in the Web of Science database, where at least one author is affiliated with the university (relative to the number of students)
Citation Rate (CR)	The average number of times that the university department's research publications (over the period 2011-2014) get cited in other research, adjusted (normalized) at the global level to take into account differences in publication years
Proportion of Top Cited Publications (PTCP)	The proportion of the university's research publications that, compared to other publications in the same field and in the same year, belong to the top 10% most frequently cited
Number of Patents Awarded (NPA)	The number of patents assigned to (inventors working in) the university (over the period 2002-2011)
Number of Spin-Offs (NSO)	The number of spin-offs (i.e. firms established on the basis of a formal knowledge transfer arrangement between the institution and the firm) recently created by the institution (per 1,000 fte academic staff)
Research and Knowledge Transfer Revenues (RKTR)	Research revenues and knowledge transfer revenues from private sources (incl. not-for profit organizations), excluding tuition fees. Measured in €1,000s using Purchasing Power Parities. Expressed per fte academic staff.

of Patents Awarded (NPA), Number of Spin-Offs (NSO) and Research and Knowledge Transfer Revenues (RKTR) are all elementary criteria descending from Knowledge Transfer.

Let us suppose that the following information has been provided, as specified in Step 1 of the scheme presented in Section 4:

1.1 The indifference, preference and veto thresholds are equal to 1, 2 and 4 respectively, for all elementary criteria; that is $q_t = 1$, $p_t = 2$ and $v_t = 4$ for all $t \in EL$.

1.2 The following interactions and antagonistic effects exist:

- CR and PTCP present a mutual-weakening effect,
- NRP and RKTR present a mutual-strengthening effect,
- PTCP presents an antagonistic effect over MGR.

1.3 The imprecise SRF method is applied to the direct subcriteria of each non-elementary criterion as follows:

- with respect to the first level criteria, TL is less important than KT that, in turn, is less important than R . The number of blank cards inserted between KT and TL belongs to the interval $[0, 1]$, while the number of blank cards inserted between R and KT belongs to the interval $[1, 2]$; moreover, the ratio between the importance of R and the importance of TL belongs to the interval $[2, 3]$,
- considering the macro-criterion TL , BGR is less important than $BGOT$ and the number of blank cards inserted between them belongs to the interval $[0, 1]$. $BGOT$ is less important than MGR and the number of cards inserted between them belongs to the interval $[1, 2]$. Finally, MGR is less important than $MGOT$ and the number of blank cards inserted between them belongs to the interval $[0, 1]$. The ratio between the importance of $MGOT$ and the importance of MGR belongs to the interval $[2, 3]$,
- with respect to the macro-criterion R , NRP is less important than $PTCP$ that, in turn, is less important than CR . The number of blank cards inserted between $PTCP$ and NRP belongs to the interval $[0, 1]$, while the number of blank cards inserted between CR and $PTCP$ belongs to the interval $[1, 2]$. The ratio between the importance of CR and the importance of NRP belongs to the interval $[2, 3]$,
- considering the macro-criterion KT , NPA is less important than NSO that, in turn, is less important than $RKTR$. The number of blank cards inserted between NSO and NPA

belongs to the interval $[1, 2]$, while the number of blank cards inserted between *RKTR* and *NSO* belongs to the interval $[0, 1]$. Moreover, the ratio between the importance of *RKTR* and the importance of *NPA* belongs to the interval $[4, 5]$.

Considering the preference information provided by the DM, we solved the LP problem (15) that gave $\varepsilon^* > 0$. This means that there exists at least one compatible vector. In order to get robust recommendations, we applied the SMAA methodology on results of the hierarchical ELECTRE III method with interactions, obtained for a sample of compatible vectors. Consequently, for each non-elementary criterion, we computed the preference, indifference and incomparability relations, as well as the indices shown in Section 4. A complete data set of results can be downloaded clicking on the following link: [dataMCHPELECTRE](#).

Table 3: Frequencies of the preference, incomparability and indifference relations between the considered universities at the comprehensive level

(a) Preference														
University / University	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	U_{12}	U_{13}	U_{14}
U Bergen (U_1)	0	0	0	13.19	0	0	100	0	0	11.24	52.22	0	100	51.07
Polytech. Inst. Bragança (U_2)	87.29	0	0	100	0	0	100	0.29	0	83.66	52.44	0	100	100
Dublin Inst. Tech (U_3)	100	99.35	0	100	0	0	100	0.29	0	100	100	14.87	100	100
European U Cyprus (U_4)	0	0	0	0	0	0	87.60	0	0	0	14.95	0	45.48	26.15
U Graz (U_5)	100	99.51	10.46	100	0	0	100	10.75	0	100	100	24.91	100	100
Lappeenranta U Tech (U_6)	100	100	100	100	100	0	100	100	99.79	100	100	100	100	100
U Lisbon (U_7)	0	0	0	0	0	0	0	0	0	0	0	0	0.94	0
U Magdeburg (U_8)	99.70	89.05	0	100	0	0	100	0	0	73.0	52.22	38.74	100	99.86
U Montpellier 1 (U_9)	100	100	99.89	100	99.81	0	100	99.59	0	100	100	99.89	100	100
MEPhl Moscow (U_{10})	24.39	0	0	0	0	0	100	0	0	0	52.22	0	100	39.84
UAS Osnabrück (U_{11})	46.62	0	0	0	0	0	100	0	0	42.80	0	0	63.14	47.78
U Pisa (U_{12})	99.70	84.77	0	100	0	0	100	21.42	0	83.66	52.22	0	100	100
Semmelweis U (U_{13})	0	0	0	0	0	0	98.32	0	0	0	0	0	0	0
U Thessaly (U_{14})	19.09	0	0	52.22	0	0	100	0	0	15.61	52.22	0	87.53	0

(b) Incomparability														
University / University	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	U_{12}	U_{13}	U_{14}
U Bergen (U_1)	0	0.29	0	86.80	0	0	0	0.29	0	2.24	0.865	0.29	0	20.88
Polytech. Inst. Bragança (U_2)	0.29	0	0.29	0	0.29	0	0	10.65	0	16.33	47.56	0.35	0	0
Dublin Inst. Tech (U_3)	0	0.29	0	0	0	0	0	99.70	0	0	0	85.13	0	0
European U Cyprus (U_4)	86.80	0	0	0	0	0	12.39	0	0	100	85.04	0	54.51	21.62
U Graz (U_5)	0	0.29	0	0	0	0	0	89.24	0	0	0	75.08	0	0
Lappeenranta U Tech (U_6)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U Lisbon (U_7)	0	0	0	12.39	0	0	0	0	0	0	0	0	0	0
U Magdeburg (U_8)	0.29	10.65	99.70	0	89.24	0	0	0	0.41	26.93	47.78	0.15	0	0.14
U Montpellier 1 (U_9)	0	0	0	0	0	0	0	0.41	0	0	0	0.10	0	0
MEPhl Moscow (U_{10})	2.24	16.33	0	100	0	0	0	26.93	0	0	0	16.33	0	36.60
UAS Osnabrück (U_{11})	0.86	47.56	0	85.04	0	0	0	47.78	0	0	0	47.78	0	0
U Pisa (U_{12})	0.29	0.35	85.13	0	75.08	0	0	0.15	0.10	16.33	47.78	0	0	0
Semmelweis U (U_{13})	0	0	0	54.51	0	0	0	0	0	0	0	0	0	0
U Thessaly (U_{14})	20.88	0	0	21.62	0	0	0	0.14	0	36.60	0	0	0	0

(c) Indifference														
University / University	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	U_{12}	U_{13}	U_{14}
U Bergen (U_1)	0	12.42	0	0	0	0	0	0	0	62.12	0.29	0	0	8.96
Polytech. Inst. Bragança (U_2)	12.42	0	0.35	0	0.20	0	0	0	0	0	0	14.87	0	0
Dublin Inst. Tech (U_3)	0	0.35	0	0	89.54	0	0	0	0.11	0	0	0	0	0
European U Cyprus (U_4)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U Graz (U_5)	0	0.20	89.54	0	0	0	0	0	0.19	0	0	0	0	0
Lappeenranta U Tech (U_6)	0	0	0	0	0	0	0	0	0.21	0	0	0	0	0
U Lisbon (U_7)	0	0	0	0	0	0	0	0	0	0	0	0	0.74	0
U Magdeburg (U_8)	0	0	0	0	0	0	0	0	0	0	0	39.67	0	0
U Montpellier 1 (U_9)	0	0	0.11	0	0.19	0.21	0	0	0	0	0	0	0	0
MEPhl Moscow (U_{10})	62.12	0	0	0	0	0	0	0	0	0	4.98	0	0	7.94
UAS Osnabrück (U_{11})	0.29	0	0	0	0	0	0	0	0	4.98	0	0	36.86	0
U Pisa (U_{12})	0	14.87	0	0	0	0	0	39.67	0	0	0	0	0	0
Semmelweis U (U_{13})	0	0	0	0	0	0	0.74	0	0	0	36.86	0	0	12.47
U Thessaly (U_{14})	8.96	0	0	0	0	0	0	0	0	7.94	0	0	12.47	0

When interpreting the results, we will call the universities by their location name. As one can see from Tables 3(a) and 4, Lappeenranta can be seen as the best among the 14 universities since it is globally preferred to all the other universities with a frequency equal to at least 99.79%. It is always

Table 4: Mean number of alternatives that are weakly preferred to each university, and mean number of alternatives to which each university is weakly preferred at the comprehensive level

University	$Mean_0^{U \text{ is } WP \text{ to}}$	$Mean_0^{WP \text{ to } U}$
U Bergen (U_1)	4.12	8.61
Polytech. Inst. Bragança (U_2)	6.52	6.01
Dublin Inst. Tech (U_3)	9.05	3.00
European U Cyprus (U_4)	1.74	7.65
U Graz (U_5)	9.36	2.90
Lappeenranta U Tech (U_6)	13.00	0.00
U Lisbon (U_7)	0.02	12.87
U Magdeburg (U_8)	7.92	2.72
U Montpellier 1 (U_9)	12.00	1.00
MEPhI Moscow (U_{10})	3.91	7.85
UAS Osnabrück (U_{11})	3.42	7.71
U Pisa (U_{12})	7.96	3.33
Semmelweis U (U_{13})	1.48	11.47
U Thessaly (U_{14})	3.56	8.94

weakly preferred to all the other universities, while any university is never weakly preferred to it. More in detail, Lappeenranta is always preferred to all alternatives except Montpellier to which it is preferred with frequency of 99.79%. In the remaining cases (0.21%) Lappeenranta is indifferent to Montpellier. As a third option after Lappeenranta and Montpellier, one could consider Graz since, on average, it is weakly preferred to 9.36 universities and only 2.90 universities are weakly preferred to it. Indeed, Lappeenranta and Montpellier are the only universities that can be preferred to Graz. At the same time, Lisbon can be considered as the worst university at the comprehensive level since it is never preferred to any other university, apart from Semmelweis, but with a very small frequency (0.94%). Moreover, looking at Table 4, one can see that Lisbon is weakly preferred to an average number of 0.02 universities, while 12.87 universities are, in average, weakly preferred to it.

Since many vectors of weights and interaction coefficients are compatible with the preference information provided by the DM, we computed, component by component, their barycenter as their mean value (see Table 5). One can observe that criterion CR is the most important among the elementary criteria, which is concordant with the requirement that its macro-criterion R is the most important at the first level. Moreover, one can observe a great importance obtained by the mutual-strengthening effect between NRP and $RKTR$, since it is equal to 0.526; this value is taken into account in computing the concordance index for pairs of alternatives at the comprehensive level.

Table 5: Barycenter of all sampled compatible vectors

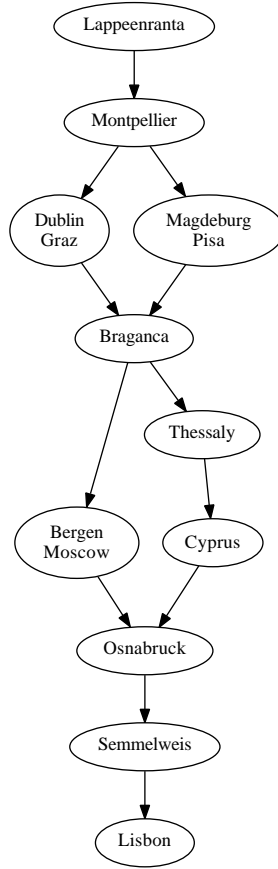
W_{TL}				W_R			W_{KT}					
0.214				0.471			0.315					
w_{BGR}	w_{MGR}	w_{BGOT}	w_{MGOT}	w_{NRP}	w_{CR}	w_{PTCP}	w_{NPA}	w_{NSO}	w_{RKTR}	$w_{CR-PTCP}$	$w_{NRP-RKTR}$	$w_{MGR-PTCP}$
0.030	0.062	0.042	0.075	0.098	0.230	0.149	0.037	0.121	0.156	-0.040	0.526	0.018

The recommended partial preorder at comprehensive level obtained for the barycenter values is shown in Figure 4. One can see that the conclusions drawn above on the base of the frequencies of preference, indifference and incomparability are confirmed by the partial preorder obtained with the weights and interaction coefficients equal to their barycenter. Lappeenranta and Montpellier take the first two positions in the ranking, while the third place is taken by Magdeburg, Pisa, Dublin and Graz. In particular, Magdeburg is indifferent to Pisa and Dublin is indifferent to Graz. Lisbon takes the last position, while Semmelweis is at the last but one position.

A similar analysis can be performed for each of the three macro-criteria. Here we shall present the results obtained for the TL macro-criterion.

Looking at Table 6(a) one can conclude that Montpellier is the best among the 14 universities on TL since it is preferred to all the other universities with a frequency equal to at least 70.91%. In

Figure 4: The partial preorder obtained at the comprehensive level by applying the hierarchical ELECTRE III method with interactions for the weights and the interaction coefficients in the barycenter



particular, Montpellier is preferred to Dublin, Lappeenranta, Osnabruck and Thessaly with the same frequency (70.91%) and only Dublin and Lappeenranta can be preferred to Montpellier. As to the second best university on TL, according to the results shown in Table 7, this is one of Lappeenranta and Dublin. Indeed, on the one hand, they are weakly preferred to an average of 10.31 and 9.61 universities, respectively, and on the other hand, an average number of only 1.15, 1.84 universities is weakly preferred to them, respectively. Comparing these two universities pairwise one can observe that Dublin is preferred to Lappeenranta with a frequency of 42.92%, while the opposite is true with a frequency of 29.08%. In the remaining 27.99% of the cases, the two universities are incomparable.

Looking at the worst university on TL, the choice goes to Bragança since it can be preferred only to Graz, while all the 13 universities are preferred to it with a frequency equal to at least 70.91%.

Regarding the other universities, some indifferences and incomparabilities can be observed (see Tables 6(c) and 6(b)). For example, Graz and Cyprus are indifferent in 70.91% of the cases, while Moscow and Thessaly are incomparable with the same frequency.

The final partial preorder obtained for TL and the barycenter values is shown in Figure 5.

Again, the partial preorder shown in Figure 5 confirms the indications coming from the frequencies of preference, incomparability and indifference, as well as from the average number of universities being weakly preferred to each university and from the mean number of universities to which each university is weakly preferred. Indeed, Montpellier is the best university on TL , followed by Dublin and Moscow. The last position is taken by Bragança, while the last but one by Lisbon.

For the sake of completeness we show in Figures 6(a) and 6(b) the partial preorders obtained for the macro-criteria R and KT obtained with the weights and interaction coefficients in the barycenter.

Analyzing more in detail the two partial preorders one can observe that:

Table 6: Frequencies of the preference, incomparability and indifference relations between the considered universities with respect to subcriterion TL

(a) Preference														
University / University	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	U_{12}	U_{13}	U_{14}
U Bergen (U_1)	0	70.91	27.99	100	100	0.13	70.91	27.99	0	20.59	27.99	27.99	27.99	27.99
Polytech. Inst. Bragança (U_2)	0	0	0	0	29.08	0	0	0	0	0	0	0	0	0
Dublin Inst. Tech (U_3)	72	100	0	100	100	42.92	100	72	29.08	29.08	72	72	72	72
European U Cyprus (U_4)	0	70.91	0	0	29.08	0	70.91	0	0	0	0	0	0	0
U Graz (U_5)	0	70.91	0	0	0	0	70.91	0	0	0	0	0	0	0
Lappeenranta U Tech (U_6)	72	100	29.08	100	100	0	100	99.86	29.08	29.08	99.86	99.86	72	99.86
U Lisbon (U_7)	0	100	0	29.08	29.08	0	0	0	0	0	0	0	29.08	0
U Magdeburg (U_8)	29.08	100	0	100	100	0	70.91	0	0	0	0	0	0	0
U Montpellier 1 (U_9)	100	100	70.91	100	100	70.91	100	100	0	100	70.91	100	100	70.91
MEPhI Moscow (U_{10})	29.08	100	27.99	100	100	0	100	29.08	0	0	0	29.08	100	0
UAS Osnabrück (U_{11})	72	100	0	100	100	0	100	72	0	29.08	0	72	29.08	72
U Pisa (U_{12})	29.08	100	0	100	100	0	70.91	0	0	0	0	0	0	0
Semmelweis U (U_{13})	0	100	0	100	100	0	70.91	0	0	0	0	0	0	0
U Thessaly (U_{14})	29.08	100	0	100	100	0	100	72	0	29.08	0	72	29.08	0

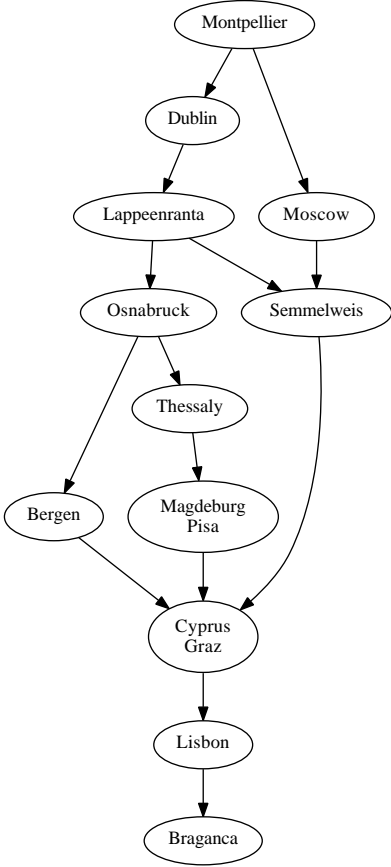
(b) Incomparability														
University / University	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	U_{12}	U_{13}	U_{14}
U Bergen (U_1)	0	29.08	0	0	0	27.85	29.08	42.92	0	50.32	0	42.92	72	42.92
Polytech. Inst. Bragança (U_2)	29.08	0	0	29.08	0	0	0	0	0	0	0	0	0	0
Dublin Inst. Tech (U_3)	0	0	0	0	0	27.99	0	27.99	0	42.92	27.99	27.99	0	27.99
European U Cyprus (U_4)	0	29.08	0	0	0	0	0	0	0	0	0	0	0	0
U Graz (U_5)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lappeenranta U Tech (U_6)	27.85	0	27.99	0	0	0	0	0	0	70.91	0	0	27.99	0
U Lisbon (U_7)	29.08	0	0	0	0	0	0	29.08	0	0	0	29.08	0	0
U Magdeburg (U_8)	42.92	0	27.99	0	0	0	29.08	0	0	70.91	0	0	100	0
U Montpellier 1 (U_9)	0	0	0	0	0	0	0	0	0	29.08	0	0	0	29.08
MEPhI Moscow (U_{10})	50.32	0	42.92	0	0	70.91	0	70.91	0	0	70.91	70.91	0	70.91
UAS Osnabrück (U_{11})	0	0	27.99	0	0	0	0	29.08	70.91	0	0	0	70.91	0
U Pisa (U_{12})	42.92	0	27.99	0	0	0	29.08	0	0	70.91	0	0	100	0
Semmelweis U (U_{13})	72	0	0	0	0	27.99	0	100	0	0	70.91	100	0	70.91
U Thessaly (U_{14})	42.92	0	27.99	0	0	0	0	29.08	70.91	0	0	0	70.91	0

(c) Indifference														
University / University	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	U_{12}	U_{13}	U_{14}
U Bergen (U_1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Polytech. Inst. Bragança (U_2)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dublin Inst. Tech (U_3)	0	0	0	0	0	0	0	0	0	0	0	0	27.99	0
European U Cyprus (U_4)	0	0	0	0	70.91	0	0	0	0	0	0	0	0	0
U Graz (U_5)	0	0	0	70.91	0	0	0	0	0	0	0	0	0	0
Lappeenranta U Tech (U_6)	0	0	0	0	0	0	0	0.13	0	0	0.13	0.13	0	0.13
U Lisbon (U_7)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U Magdeburg (U_8)	0	0	0	0	0	0.13	0	0	0	0	27.99	100	0	27.99
U Montpellier 1 (U_9)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MEPhI Moscow (U_{10})	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UAS Osnabrück (U_{11})	0	0	0	0	0	0.13	0	27.99	0	0	0	27.99	0	27.99
U Pisa (U_{12})	0	0	0	0	0	0.13	0	100	0	0	27.99	0	0	27.99
Semmelweis U (U_{13})	0	0	27.99	0	0	0	0	0	0	0	0	0	0	0
U Thessaly (U_{14})	0	0	0	0	0	0.13	0	27.99	0	0	27.99	27.99	0	0

Table 7: Mean number of alternatives that are weakly preferred to each university, and mean number of alternatives to which each university is weakly preferred with respect to subcriterion TL

University	$Mean_{TL}^{U \text{ is WP to}}$	$Mean_{TL}^{WP \text{ to } U}$
U Bergen (U_1)	5.31	4.32
Polytech. Inst. Bragança (U_2)	0.29	12.13
Dublin Inst. Tech (U_3)	9.61	1.84
European U Cyprus (U_4)	2.42	11.00
U Graz (U_5)	2.13	11.58
Lappeenranta U Tech (U_6)	10.31	1.15
U Lisbon (U_7)	1.87	10.25
U Magdeburg (U_8)	5.56	6.29
U Montpellier 1 (U_9)	11.84	0.58
MEPhI Moscow (U_{10})	6.15	2.37
UAS Osnabrück (U_{11})	8.30	3.55
U Pisa (U_{12})	5.56	6.29
Semmelweis U (U_{13})	3.99	4.87
U Thessaly (U_{14})	7.15	4.27

Figure 5: The partial preorder obtained on TL by applying the hierarchical ELECTRE III method with interactions for the weights and the interaction coefficients in the barycenter



- with respect to R , Bergen and Lisbon are equally the best and they are followed by Graz and Montpellier. Looking at the tail of the ranking, Moscow and Thessaly take the last two positions in this order. Some incomparabilities exist such that one between Bragança and Semmelweis or that one between Cyprus and Osnabrück,
- with respect to KT , Lappeenranta is the best among the considered universities, followed by Bragança and Montpellier. Looking at the tail of the ranking, Lisbon and Semmelweis are equally the worst, while Bergen takes the last but one position.

As underlined in point 5.3 of the scheme presented in the previous section, let us observe that when applying the ELECTRE III method with interactions, we obtained different final partial preorders at comprehensive level as well as at partial level even if we have shown here only the final partial preorders obtained considering the barycenter of the sampled weights and interaction coefficients. In particular, we obtained 39 partial preorders at comprehensive level, 5 partial preorders on Teaching and Learning, and 1 partial preorder only on Research and Knowledge Transfer, being therefore the two partial preorders shown in Figures 6(a) and 6(b), respectively. The frequencies of the obtained partial preorders as well as the barycenter of the weights for which the same partial preorders are obtained can be downloaded clicking on the following link: [dataMCHPELECTRE](#).

To conclude this section, we would like to underline the usefulness of taking into account the MCHP in the decision problem at hand and, in general, in all decision problems involving hierarchical criteria. Let us consider universities of Moscow and Osnabrück. Looking at Table 1, one can observe that the two universities do not dominate each other at any of the three considered macro-criteria. Consequently, looking only at the performances of the universities on the elementary criteria nothing

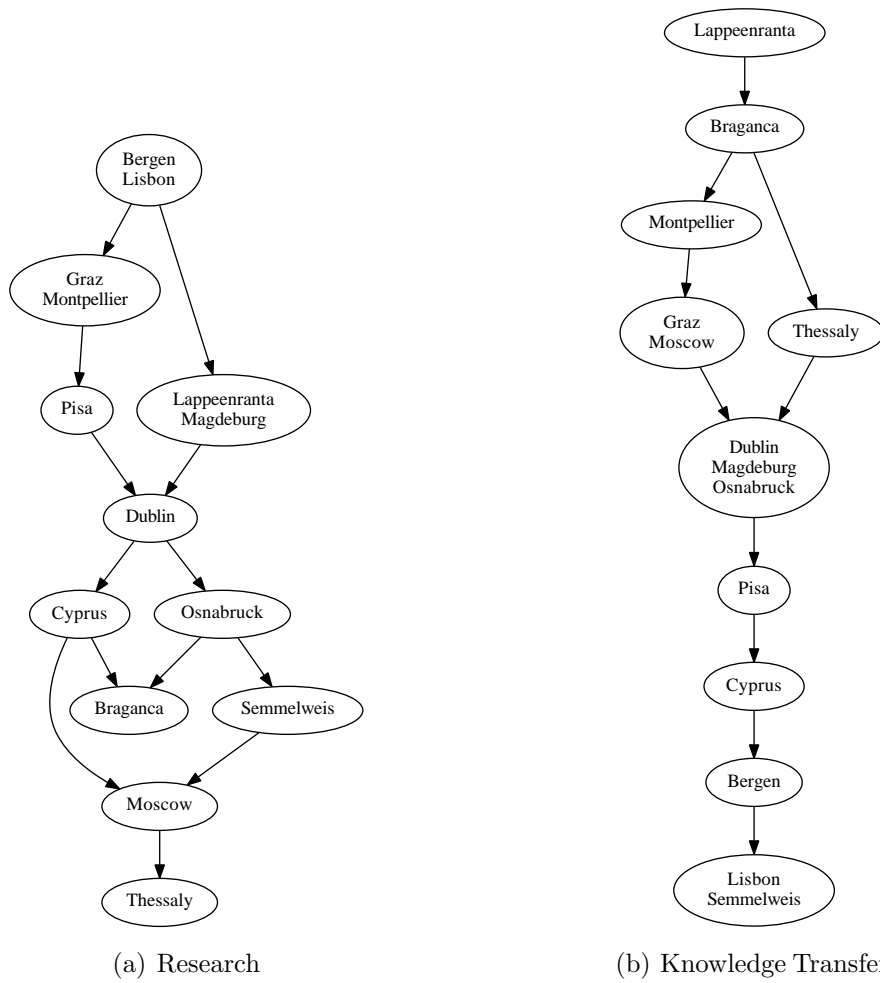
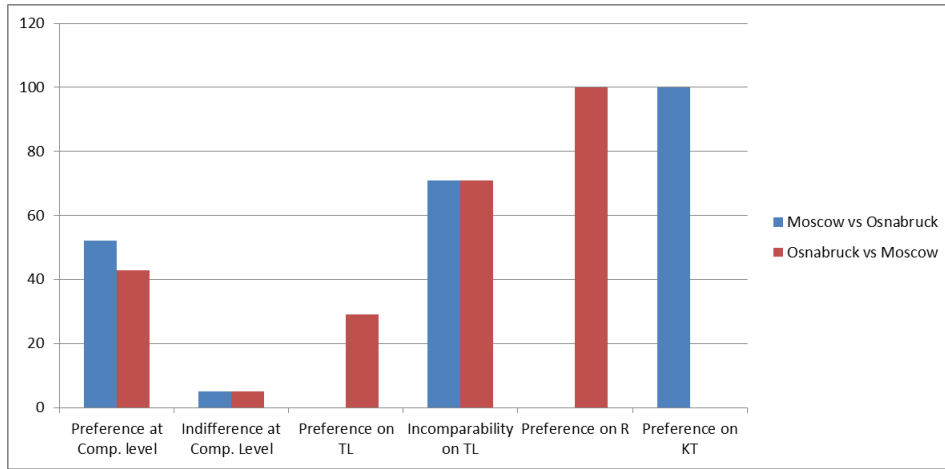


Figure 6: The partial preorders obtained on the macro-criteria R and KT by applying the proposed methodology for the weights and interaction coefficients in the barycenter

Figure 7: Comparison between the universities of Moscow and Osnabrück considering the frequencies of the preference, indifference and incomparability between them at the comprehensive level, as well as at the three macro-criteria



could be said about the preference of one university over the other, neither at the comprehensive level, nor at the level of macro-criteria. Using the preference information provided by the DM, one can instead observe the positive frequencies shown in Figure 7. Moscow is preferred to Osnabrück slightly more frequently than vice versa at the comprehensive level, while they are indifferent on the remaining cases. Going more in depth with our analysis, some interesting results can be noted. While on TL they are incomparable with a frequency of 70.91%, and Osnabrück is preferred to Moscow in 29.09% of the cases, two opposite situations can be observed for the macro-criteria R and KT . On R , Osnabrück is always preferred to Moscow, while an inverse preference is always true on KT . We think that even if no conclusion can be drawn for these universities at the comprehensive level, the frequencies of preference, indifference and incomparability relations between them at the level of macro-criteria can be useful for the DM.

6. Conclusions

In this paper, we proposed a methodology of multiple criteria ranking that aims to answer the current challenges of scientific decision aiding. These are:

- (a) Handling of multi-faceted performances through consideration of many criteria structured hierarchically.
- (b) Consideration of interaction between criteria that may show up as mutual-weakening effects, mutual-strengthening effects, and antagonistic effects.
- (c) Acceptance of intransitivity or incompleteness in the recommended ranking of the alternatives, due to preference modeling by an outranking relation.
- (d) Consideration of robustness concerns related to non-univocal values of preference model parameters through a stochastic analysis of recommendations.

The proposed methodology handles the above issues jointly, for the first time in the literature.

The reader has certainly noticed that elaboration of this comprehensive methodology required more than a simple gathering of the existing methods coping with each particular issue. As they are interrelated, a careful adaptation of the methods contributing to the comprehensive methodology was necessary. In particular, we combined for the first time the ELECTRE III with the Multiple Criteria Hierarchy Process (MCHP) to handle all types of interactions between criteria structured in a

hierarchical way. We also involved in MCHP an imprecise elicitation of criteria weights, generalizing the SRF method. In order to explore the plurality of rankings obtained by the ELECTRE III method for possible sets of criteria weights, we applied the SMAA method that permits to draw robust conclusions in terms of rankings and preference relations at each level of the hierarchy of criteria.

The choice of the ELECTRE III method as a central tool of our methodology is motivated by its capability to handle an imperfect knowledge encountered in real-world decision aiding situations. This imperfect knowledge comes from several sources [24]: uncertainty, imprecision, ill-determination, and an inherent arbitrariness when building a family of criteria. To cope with these imperfections, the basic ELECTRE methods assume that criteria are equipped with indifference and preference thresholds, serving as discrimination thresholds in pairwise comparisons of alternatives on particular criteria. Criteria equipped with these thresholds are called pseudo-criteria. Moreover, the ELECTRE methods construct an outranking relation as a preference model that permits to represent not only preference and indifference but also incomparability in the set of alternatives, due to the above mentioned imperfect knowledge, as well as consideration of a possible veto against a hypothesis of outranking from the part of discordant criteria. The outranking preference model is also a non-compensatory model of multiple criteria aggregation, that is important for the fact that in practice the criteria are of heterogeneous nature. The addition of interaction between criteria structured in a hierarchy tree makes of the extended ELECTRE III method a powerful tool able to cope with real-world decision problems.

The hierarchical structure of the family of criteria permits representation of preferences and checking of ranking recommendations not only with respect to the whole family of criteria (at the comprehensive level) but also with respect to any subset of criteria descending from a particular node of the hierarchy tree. This feature of the proposed methodology augments transparency in decision aiding and helps the DM to check the effects of criteria interaction on the ranking recommendations at different levels of the hierarchy.

An important component of the proposed methodology is the imprecise SRF method that helps the DM to get the weights of criteria and sub-criteria on the base of not completely determined preference information. In consequence of considering an imprecise preference information in SRF, there is an infinity of compatible vectors of weights satisfying the constraints translating this preference information. For this reason, in the last stage of the proposed methodology, we apply the SMAA method that stochastically samples the compatible vectors and draws statistical conclusions on the resulting preference relations and ranking recommendations. In this way, the robustness as conceived by Roy [25] for whom “*robustness involves concerns that must be taken into account a priori, at the time that the problem is formulated*”, is well handled.

The whole methodology has been illustrated on an example of university ranking. In this application, all steps of the methodology were described in a simple way for an easy understanding. We believe this proposal makes a large step towards more realistic and transparent multiple criteria decision aiding.

Acknowledgments

The first and the third authors wish to acknowledge the funding by the “FIR of the University of Catania BCAEA3, New developments in Multiple Criteria Decision Aiding (MCDA) and their application to territorial competitiveness”.

References

- [1] <http://www.umultirank.org>.

- [2] J. Almeida-Dias, J.R. Figueira, and B. Roy. The software ELECTRE III-IV methodology and user manual (Version 3.X). *University Paris-Dauphine Lamsade*, 2006.
- [3] J.P. Brans and Ph. Vincke. A preference ranking organisation method: The PROMETHEE method for MCDM. *Management Science*, 31(6):647–656, 1985.
- [4] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1954.
- [5] S. Corrente, S. Greco, , and R. Słowiński. Multiple criteria hierarchy process for ELECTRE TRI methods. *European Journal of Operational Research*, 252(1):191–203, 2016.
- [6] S. Corrente, S. Greco, M. Kadziński, and R. Słowiński. Robust ordinal regression in preference learning and ranking. *Machine Learning*, 93:381–422, 2013.
- [7] S. Corrente, S. Greco, and R. Słowiński. Multiple Criteria Hierarchy Process in Robust Ordinal Regression. *Decision Support Systems*, 53(3):660–674, 2012.
- [8] S. Corrente, S. Greco, and R. Słowiński. Multiple Criteria Hierarchy Process with ELECTRE and PROMETHEE. *Omega*, 41:820–846, 2013.
- [9] J.R. Figueira, S. Greco, and B. Roy. ELECTRE methods with interaction between criteria: An extension of the concordance index. *European Journal of Operational Research*, 199(2):478–495, 2009.
- [10] J.R. Figueira, S. Greco, B. Roy, and R. Słowiński. An overview of ELECTRE methods and their recent extensions. *Journal of Multicriteria Decision Analysis*, 20:61–85, 2013.
- [11] J.R. Figueira and B. Roy. Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139:317–326, 2002.
- [12] S. Greco, M. Ehrgott, and J.R. Figueira. *Multiple Criteria Decision Analysis: State of the Art Surveys*. Springer, Berlin, 2016.
- [13] S. Greco, V. Mousseau, and R. Słowiński. Ordinal regression revisited: Multiple criteria ranking using a set of additive value functions. *European Journal of Operational Research*, 191(2):416–436, 2008.
- [14] S. Greco, V. Mousseau, and R. Słowiński. Robust ordinal regression for value functions handling interacting criteria. *European Journal of Operational Research*, 239(3):711–730, 2014.
- [15] A. Hatami-Marbini and M. Tavana. An extension of the ELECTRE I method for group decision-making under a fuzzy environment. *Omega*, 39(4):373–386, 2011.
- [16] M. Kadziński and T. Tervonen. Robust multi-criteria ranking with additive value models and holistic pair-wise preference statements. *European Journal of Operational Research*, 228(1):169–180, 2013.
- [17] M. Kadziński and T. Tervonen. Stochastic ordinal regression for multiple criteria sorting problems. *Decision Support Systems*, 55(11):55–66, 2013.
- [18] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. J. Wiley, New York, 1993.
- [19] R. Lahdelma, J. Hokkanen, and P. Salminen. SMAA - stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 106(1):137–143, 1998.

- [20] P. Leskinen, J. Viitanen, A. Kangas, and J. Kangas. Alternatives to incorporate uncertainty and risk attitude in multicriteria evaluation of forest plans. *Forest Science*, 52(3):304–312, 2006.
- [21] R.D. Luce. Semiorders and a theory of utility discrimination. *Econometrica*, 24:178–191, 1956.
- [22] V. Mousseau, J.R. Figueira, L. Dias, C. Gomes da Silva, and J. Clímaco. Resolving inconsistencies among constraints on the parameters of an MCDA model. *European Journal of Operational Research*, 147(1):72–93, 2003.
- [23] B. Roy. ELECTRE III: Un algorithme de classements fondé sur une représentation floue des préférences en présence de critères multiples. *Cahiers du CERO*, 20(1):3–24, 1978.
- [24] B. Roy. Main sources of inaccurate determination, uncertainty and imprecision in decision models. *Mathematical and Computer Modelling*, 12(10-11):1245–1254, 1989.
- [25] B. Roy. Robustness in operational research and decision aiding a multi-faceted issue. *European Journal of Operational Research*, 200:629–638, 2010.
- [26] B. Roy and D. Bouyssou. *Aide Multicritère à la Décision: Méthodes et Cas*. Economica, Paris, 1993.
- [27] J. Simos. *Evaluer l’Impact sur l’Environnement: Une Approche Originale par l’Analyse Multi-critère et la Négociation*. Presses Polytechniques et Universitaires Romandes, Lausanne, 1990.
- [28] J. Simos. *L’Evaluation Environnementale: Un Processus Cognitif Négocié*. Ph.D. Thesis. DGF-EPFL, Lausanne, 1990.
- [29] E. Siskos and N. Tsotsolas. Elicitation of criteria importance weights through the Simos method: A robustness concern. *European Journal of Operational Research*, 246(2):543–553, 2015.
- [30] R.L. Smith. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32:1296–1308, 1984.
- [31] M. Sugeno. *Theory of Fuzzy Integrals and its Applications*. Tokyo Institute of Technology, 1974.
- [32] T. Tervonen and J.R. Figueira. A survey on stochastic multicriteria acceptability analysis methods. *Journal of Multi-Criteria Decision Analysis*, 15(1-2):1–14, 2008.
- [33] T. Tervonen, H. Hakonen, and R. Lahdelma. Elevator planning with stochastic multicriteria acceptability analysis. *Omega*, 36(3):352–362, 2008.
- [34] T. Tervonen, G. Van Valkenhoef, N. Bastürk, and D. Postmus. Hit-and-run enables efficient weight generation for simulation-based multiple criteria decision analysis. *European Journal of Operational Research*, 224:552–559, 2013.
- [35] G. Van Valkenhoef, T. Tervonen, and D. Postmus. Notes on “Hit-And-Run enables efficient weight generation for simulation-based multiple criteria decision analysis”. *European Journal of Operational Research*, 239(3):865–867, 2014.